

**Shared representations in language processing and verbal
short-term memory: The case of grammatical gender**

Dissertation

zur Erlangung des akademischen Grades
eines Doktors der Philosophie
der Philosophischen Fakultäten
der Universität des Saarlandes

vorgelegt von

Judith Schweppe, M.A.

aus Bünde

Saarbrücken, 2006

Der Dekan:

Prof. Dr. Rainer Krause

Berichterstatter:

PD Dr. Ralf Rummer

Prof. Dr. Herbert J. Schriefers

Prof. Dr. Dirk Wentura

Prof. Dr. Werner H. Tack

Tag der Disputation:

09.11.2006

Acknowledgements

There are many people that I want to thank for their contributions to this work. Some of them shall be mentioned in the following:

First of all, I want to thank Ralf Rummer, who was my advisor and mentor throughout this project. He provided a lot of “resources” without which this thesis definitely would not have come about, among these are theoretical input, fruitful discussions, and infectious enthusiasm.

I am also very grateful to Herbert Schriefers, Dirk Wentura, and Werner Tack, who have “volunteered” to review my thesis. Thanks in advance for your time and effort!

In addition, I want to thank Christoph Scheepers, Dirk Wentura, and Markus Pospeschill for helpful discussion concerning the data analysis and Kay Bock, Randi Martin, and two anonymous reviewers for their constructive comments on an article, in which part of this work has been reported.

I am also very grateful to Alissa, Sonia, and Ursel, who have proofread my thesis and have tried to polish my English. Thanks a lot for your helpful comments.

This research was supported by a grant from the German Research Foundation to Ralf Rummer (DFG; by grant Ru 891/1-1 and Ru 891/1-2).

I owe many thanks to our student researchers, who helped me with data collection and data analysis. In this context, I especially want to thank Anne, who conducted Experiment 5 as her diploma thesis. Thank you, Anja, Anne, Beata, Franca, Isabel, Jasmin, Jasmin, Jenny, Melanie, Ralph, and Vera!

Further thanks go to all members of “The Arbeitseinheit” and in particular to my fellow PhD-students. Our professional and private workshops were always quite inspiring.

For their “social help”, I want to thank all those that have accompanied me to the Mensa within the last years. In changing cast, these are Anne, Daniela, Franca, Frank, Heike, Jasmin, Jenny, Julia, Juliane, Kerstin, Kristina, Marc, Nadine, Rainer, Ralf, Ralph, and Rebecca. Thanks for numerous conversations about BILD, broccoli, death, desserts, Gerhard Schröder, goulash, Mensa improvements, the ongoing construction works, sexual equality, soccer, trade unions, and many essential things more.

And Arne, I want to thank you for your love and support. Your cooking, care, and company brought me through this thesis, particularly through the last [REDACTED] weeks.

My family and friends (many of whom have already been mentioned above): Thanks for all! I hope you’ll always be around me! (In particular, I want to thank my mum for forgiving me that I did not become an Olympic Champion – and did not even come close to that.)

Last but not least, I want to thank in advance all those that will celebrate with me.

Or, to say it with Andi Brehme and Horst Hrubesch: *Ich sage nur ein Wort: Vielen Dank!*

Table of Contents

List of Tables.....	v
List of Figures.....	vii
0. Introduction.....	1
1. Verbal working memory	7
1.1 <i>Baddeley's model of working memory and the storage of</i> <i>verbal information</i>	8
1.2 <i>Working memory as the activated part of long-term memory</i>	14
1.3 <i>Language-based models of short-term memory</i>	16
1.4 <i>Short-term memory for sentences</i>	21
1.4.1 <i>Memory for surface and gist information as a function</i> <i>of a sentence's state of processing</i>	23
1.4.2 <i>The conceptual regeneration hypothesis</i>	28
1.4.3 <i>Memory for sentences as a function of sentence processing</i>	30
2. Grammatical gender.....	43
2.1 <i>Grammatical gender in language comprehension</i>	44
2.2 <i>Grammatical gender in language production</i>	47
3. Objectives of the present study	53
4. Experiment 1	61
4.1 <i>Method</i>	62
4.2 <i>Results</i>	66
4.3 <i>Discussion</i>	71
5. Experiment 2.....	73
5.1 <i>Method</i>	73
5.2 <i>Results</i>	74
5.3 <i>Discussion</i>	78
6. Experiment 3.....	81
6.1 <i>Method</i>	81
6.2 <i>Results</i>	82
6.3 <i>Discussion</i>	86
7. Experiment 4.....	89
7.1 <i>Method</i>	89
7.2 <i>Results</i>	92
7.3 <i>Discussion</i>	97
7.4 <i>Joint Discussion of Experiments 1-4</i>	98
8. Experiment 5.....	103
8.1 <i>Method</i>	105
8.2 <i>Results</i>	109

8.2.1 Lure Intrusions.....	110
8.2.2 Clause Recall.....	118
8.3 Discussion.....	121
9. General Discussion.....	127
9.1 Implications for models of short-term memory.....	128
9.2 Implications for models of language production and language comprehension.....	135
9.3 Final words and future work.....	141
References.....	149
Appendix.....	163
A Materials.....	163
A 1 German originals.....	163
A 2 English glosses.....	178
B Means and standard deviations for participants' overall recall accuracy in Experiments 1 to 4.....	189
C Results of Experiment 5 with all 30 items.....	189
C 1 Table of results.....	189
C 2 Lure Intrusions.....	190
C 3 Recall accuracy.....	192
D Instructions.....	193
D 1 Instructions for the experiments: German originals.....	193
D 2 English glosses.....	195
D 3 Instructions pretests: German originals.....	197
D 4 Instructions pretests: English glosses.....	198
German Summary.....	199

List of Tables

Table 1: Means and standard deviations for the target and lure words in the pretests for Experiment 1.....	63
Table 2: Absolute frequencies and percentages of gender congruent and gender incongruent lure intrusions, correctly reproduced target words, and target omissions as a function of lure condition in Experiment 1.....	66
Table 3: Absolute frequencies and percentages of gender congruent and gender incongruent lure intrusions, correctly reproduced target words, and target omissions as a function of lure condition in Experiment 2.....	75
Table 4: Absolute frequencies and percentages of gender congruent and gender incongruent lure intrusions, correctly reproduced target words, and target omissions as a function of lure condition in Experiment 3.....	83
Table 5: Means and standard deviations for the target and lure words in the pretests for Experiment 4.....	91
Table 6: Absolute frequencies and percentages of gender congruent and gender incongruent lure intrusions, correctly reproduced target words, and target omissions as a function of lure condition and morphophonological specification for gender in Experiment 4.....	93
Table 7: Means and standard deviations for the target and lure words in the pretests for Experiment 5.....	107
Table 8: Absolute frequencies and percentages of gender congruent and gender incongruent lure intrusions, correctly reproduced target words, and target omissions as a function of lure condition and state of processing in Experiment 5.....	112
Table 9: Absolute frequencies and percentages of gender congruent and gender incongruent lure intrusions, correctly reproduced target words, and target omissions as a function of lure condition and state of processing in Experiment 5 (with all 30 items included)	189
Table 10: Proportion of correctly recalled words per clause as a function of state of processing.....	192
Table 11: Proportion of perfectly recalled clauses as a function of state of processing	192

List of Figures

Figure 1: The phonological loop in Baddeley's (1986) model of working memory.	10
Figure 2: Baddeley's (2000) working memory model and its relation to long-term memory.	12
Figure 3: Cowan's (2001) tripartite memory model.	15
Figure 4: Martin et al.'s (1999) conception of verbal short-term memory.	17
Figure 5: Percentage of correct judgments for each probe type as a function of the number of syllables between original and probe sentence in Sachs' (1967) experiment.	24
Figure 6: Accuracy in judging changes for each probe type as a function of presentation modality and of the number of syllables between original and probe sentence (Sachs, 1974).	25
Figure 7: Proportion of correctly recalled words per clause as a function of syntactic structure in Jarvella (1972).	26
Figure 8: Representations involved in sentence regeneration (adapted from Rummer & Engelkamp, in press).	31
Figure 9: Percentages of gender congruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 1.	69
Figure 10: Percentages of gender incongruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 1.	70
Figure 11: Percentages of induced lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 1.	71
Figure 12: Percentages of gender congruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 2.	76
Figure 13: Percentages of gender incongruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 2.	77
Figure 14: Percentages of induced lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 2.	78
Figure 15: Percentages of gender congruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 3.	84
Figure 16: Percentages of gender incongruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 3.	85

Figure 17: Percentages of induced lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 3.....	86
Figure 18: Percentages of gender congruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 4.	95
Figure 19: Percentages of gender incongruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 4.	95
Figure 20: Percentages of induced lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 4.....	96
Figure 21: Percentages of gender congruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 5.	114
Figure 22: Percentages of gender incongruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 5.	115
Figure 23: Percentages of induced lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 5.....	116
Figure 24: Proportion of correctly reproduced words per clause as a function of syntactic structure in Experiment 5.....	119
Figure 25: Proportion of perfectly reproduced clauses as a function of syntactic structure in Experiment 5.	120

Da hab ich gedacht, ich tu ihn ihm rein in ihn ihm sein Tor.

(Then I thought, I put it him in into him him his goal.)

Horst Hrubesch, German soccer player

0. Introduction

This thesis addresses a question in the context of verbal working memory. With respect to everyday activities, the ability to remember verbal information over a short period of time has stereotypically been associated with remembering telephone numbers. By now, nearly everybody has a telephone which can save more numbers than they can have friends so that this ability seems dispensable. However, that we are able to briefly remember verbal information is helpful also in other cases. For instance, for calculating you often have to store by-products until you get a final result. You need to remember what you wanted to buy until you have reached the supermarket. Or for understanding the punch line of a joke it is often necessary to exactly remember previous parts of the joke because one has to revise one's initial understanding to capture what is funny. When you are engaged in a dialogue, it is also necessary to keep a kind of discourse protocol to keep track of what your partner has said so far and what you have said so far.

Verbal working memory is often primarily associated with a phonological representation that is retained over a short period of time. However, verbal materials carry a lot more information (e.g., meaning information) and in the last years models have come about that emphasize the need to incorporate a wider range of linguistic representations into the concept of verbal working memory. Such models are the theoretical grounds of my study. Specifically, they assume that verbal short-term memory and language processing rely on the *same representations*.

These models are termed *language-based models* of short-term memory (see e.g., Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Martin & Gupta, 2004; Martin & Saffran, 1997; Martin & Freedman, 2001; Martin, Lesch, & Bartha, 1999). One

kind of evidence for their assumption that short-term retention of verbal materials is closely linked to linguistic representations and to the architecture of the language processing system is the finding that lexical-semantic features influence recall performance. Among these factors are concreteness or how often one normally encounters a certain word (i.e., word frequency). Language-based approaches thus question former models which have viewed verbal short-term memory predominantly as a phonological store operating independently of language-related long-term knowledge and they emphasize the need to systematically investigate parallels between perception, production, and retention of words (in isolation or on lists).

In the focus of the present study is another model that assumes that representations are shared in language processing and verbal short-term memory (Rummer & Schweppe, in prep.). It postulates strong interdependence between processing and retention of verbal materials. The assumed interdependence is such that what types of information are retained and how long they are retained depends on what is *needed* for subsequent language comprehension or subsequent language production. For instance, one needs to know the meaning of previous parts of an utterance to understand what follows. Or the number of a noun (i.e., either singular or plural) has to be retained until the respective verb is produced so that noun and verb agree in number (i.e., so that both noun and verb are in singular form or so that both noun and verb are in plural form).

In the course of my study, I want to investigate predictions that follow from language-based models in general and from the last-mentioned model in particular. Although most studies that have addressed parallels between language processing and short-term memory so far have concentrated on retention of single words or lists of words (such as a shopping list), this prediction should also be appropriate beyond the single word level. Both what has to be processed and what has to be stored differs between telephone numbers, mathematical terms, shopping lists, and texts or dialogues. Following from the assumptions that processing and retention of verbal information involve the same representations, it can also be assumed that remembering single or unrelated words differs from remembering words that are related to each other syntactically and conceptually as is the case in sentences or texts. The model by Rummer and Schweppe (in prep.) wants to provide a more general basis for the lan-

guage dependency of verbal short-term memory and, as a consequence, addresses the retention of syntactically structured materials.

Hence, my experiments with German-speaking participants focus on short-term retention of sentences and short texts. I want to test whether a type of information is retained in short-term memory that has not been considered relevant for memory tasks so far: morphosyntactic information. If one could indeed demonstrate a recall contribution of morphosyntactic information – a type of information that is usually attributed to language processing and not to working memory – this would make a strong case for the idea that representations and processes are shared between language processing and verbal short-term memory.

What is more, the role of morphosyntactic information in short-term memory is important in the context of the model on which this study is based. As indicated above, this account does not only assume that processing and retention of language make use of the same representations but that it is indeed language processing that *drives* short-term memory. This means that those pieces of information are maintained that are often needed for understanding or producing subsequent utterances. In addition to a representation of meaning, in German this includes, for instance, the grammatical gender of the noun phrases that have occurred so far.

In German, grammatical gender is a feature that is often needed for the interpretation of personal pronouns. Personal pronouns have hardly any meaning independent of the noun they refer to. They are specified with respect to case, number, and gender and have to agree with their referents in number and gender. The following example (1) with two personal pronouns is meant to illustrate the importance of grammatical gender for pronoun interpretation by illustrating the difficulty that arises when the gender of the pronouns is non-informative.

- (1) Da hab ich gedacht, ich tu *ihn*_[accusative, singular, masculine] *ihm*_[dative, singular, masculine] rein in *ihn*_[accusative, singular, masculine] *ihm*_[dative, singular, masculine] sein_[possessive pronoun, singular, masculine] Tor.
Literal translation: Then have I thought I do *him*_[accusative, singular, masculine] *him*_[dative, singular, masculine] in into *him*_[accusative, singular, masculine] *him*_[dative, singular, masculine] his_[possessive pronoun, singular, masculine] goal.

Gloss: Then I thought I put it him in into (him him) his goal.

(quotation from Horst Hrubesch, German soccer player in the 1970s and the 1980s)

Since both the pronoun “ihn” and the pronoun “ihm” (and thus both referents) have masculine gender, grammatical gender cannot be used to identify the correct referents. Several factors make the quotation from the former German soccer player hard to understand. There is a tiny (morpho-)syntactic error in the second part of the sentence (“in sein Tor” would have made a correct noun phrase) and the phonological similarity between the personal pronouns and the preposition increases confusability. Yet there is another difficulty, which is the critical one in the present context: namely, the gender identity between the pronouns and their referents. The first masculine pronoun “ihn” refers to the ball and the second masculine pronoun “ihm” to the other team’s keeper. Reference is thus ambiguous. Had Hrubesch referred to the ball with an alternative noun with different gender, for instance with “das_[neuter] Leder” (literal translation: the leather), the sentence would have been “Da hab ich gedacht, ich tu *es*_[neuter] ihm_[masculine] rein ...” and could be understood more easily. So, particularly with pronouns, which are defined by their morphosyntactic properties, grammatical gender is an essential information type in German. If a model predicts that the relevance of a representation for (subsequent) sentence comprehension (or equivalently subsequent sentence production) determines whether a certain type of representation will indeed be maintained, this suggests that grammatical gender should play a role in recall of verbal materials. By investigating the recall contribution of grammatical gender information, the present study therefore addresses the issue of the relationship between verbal working memory and language processing.

Overview. Following this introduction, the first chapter of my thesis aims at sketching a selection of models of (verbal) working memory. In its first two subchapters, two influential models of working memory will briefly be described. The first one is the seminal multi-component model of Baddeley and Hitch (1974; further developments e.g., Baddeley, 1986, 2000), the second one is the embedded processes model by Cowan (1995, 1999, 2001). The sketch of these models is followed by an introduction to language-based models of verbal short-term memory that focus on the

relationship between word perception, word production, and word retention (Dell et al., 1997; Martin et al., 1999). The final part of the chapter on verbal working memory specifically addresses the question of short-term memory for *sentences* (and for short text passages). Among others, the language-based account of short-term memory for syntactically structured verbal material that provides the theoretical framework for this study (Rummer & Schweppe, in prep., as indicated above) is introduced here. This account predicts that grammatical gender information should contribute to short-term retention of syntactically structured materials because a representation of the grammatical gender of earlier nouns or noun phrases is needed for successful processing of sentences and texts.

To substantiate the assumption that the here-investigated morphosyntactic feature (i.e., grammatical gender) is indeed relevant in language comprehension and language production, I will review psycholinguistic studies on gender processing in Chapter 2.

Chapter 3 sums up the theoretical foundations laid out before, reports the objectives of the present study, introduces the main paradigm applied here, and gives a brief overview of my experiments.

In Chapters 4 to 8, I will report the five experiments conducted to test whether grammatical gender information makes a significant contribution to short-term recall of syntactically structured verbal materials. Such a contribution is predicted on the grounds of language-based models of short-term memory and particularly on the grounds of Rummer and Schweppe's model. The first four experiments address the question of a grammatical gender contribution to recall of sentences. Experiments 2 and 3 aim at reducing the degree to which phonological information possibly contributes to the observed effect. Experiment 4 is conducted to provide evidence against a lexical-based explanation of the effect that I attribute to grammatical gender information.¹ After a joint discussion of these four sentence recall experiments in the final part of Chapter 7, the last experiment extends the investigation of grammatical gender involvement in short-term memory to recall of short texts.

Chapter 9 discusses possible implications for the two research fields between which this study is located: models of verbal short-term memory and psycholinguistic

¹ Experiments 1, 2, and 4 have also been published in a forthcoming paper in the *Journal of Memory and Language* (Schweppe & Rummer, 2007).

tic theories (of language comprehension and language production). The thesis ends with final remarks and an outlook concerning shortcomings of this study and possibilities for future work.

1. Verbal working memory

This chapter will give an introduction to the current state of important models of (verbal) working memory. The selection of models that I will describe here is far from providing a comprehensive overview of all approaches to working memory. For instance, influential accounts focus on individual differences in working memory capacity and on the relation of working memory capacity to performance in complex cognitive tasks (e.g., Engle, Kane, & Tuholski, 1999; Just & Carpenter, 1992). Others focus on the role of experience in working memory performance (Ericsson & Kintsch, 1995; Ericsson & Delaney, 1999). However, the models I will focus on in this chapter are meant to enlighten at least some different perspectives on the relationship between working memory, language comprehension, and language production. A more comprehensive survey of accounts of working memory is given by Miyake and Shah (1999), who have assembled contributions from a wide range of working memory researchers.

The first two theories that will be sketched here (Baddeley's model in Subchapter 1.1 and Cowan's model in Subchapter 1.2) provide more general accounts of working memory and have been designed to explain short-term storage and processing of visual and verbal information. The models described in the following subchapter (1.3) are restricted to short-term memory for verbal materials and emphasize parallels between language processing and verbal short-term memory. While all of these models are predominantly concerned with retention of words or word lists, the three accounts that will be introduced in the final subchapter (1.4) address the question of memory for sentences. Despite their differences, each of these accounts posits that different processes are involved in the retention of lists and the retention of sentences.

Before sketching these models, a terminological problem or terminological uncertainty needs to be addressed. As one might have noticed, I have used both the term "verbal short-term memory" and the term "verbal working memory" without providing a definition that differentiates between these two. Although "there is not always a clear-cut distinction between working memory and the still prevalent concept of 'short-term memory'" (Shah & Miyake, 1999, p. 2), a common distinction defines short-term memory as just a storage mechanism while working memory includes storage *and* processing operations (e.g., Daneman & Carpenter, 1980, 1983;

Engle et al., 1999). The present study focuses on recall of sentences and short texts and its theoretical foundation is a model that states that recall of syntactically structured materials is a regenerative process that does per se involve more than just storage. Moreover, it conceives of short-term *storage* of verbal materials as a function of language *processing*. Inherent to this view is the assumption that storage cannot be structurally separated from processing. In consequence, this account implies that the distinction between pure storage of information and storage in combination with information processing is dispensable in the context of retention of syntactically structured materials. In other words, it implies that the distinction between verbal short-term memory and verbal working memory is dispensable in the context of the present investigation. For that reason, I will not distinguish between these terms and will use them interchangeably.

1.1 Baddeley's model of working memory and the storage of verbal information

Probably the most influential model of working memory was proposed by Alan Baddeley and Graham Hitch (1974, for a comprehensive overview cf. Baddeley, 1986). In this model, that has been developed further by Baddeley (see 2000 for the most recent version), working memory is divided into three components. Baddeley and Hitch (1974) assumed two slave systems, one specified for storage and processing of verbal and acoustic materials (in earlier versions termed the *articulatory loop*, later termed the *phonological loop*) and one for visual and spatial information (the *visuospatial sketchpad*). Both of these slave systems consist of two parts, a passive storage component and an active rehearsal mechanism which serves to refresh the stored information in order to prevent it from decay. The phonological loop and the visuospatial sketchpad are supervised by an attentional system (the *central executive*). The central executive allocates resources to the slave systems (for a discussion of possible, separable functions of the central executive see e.g., Baddeley, 1993; Miyake, Friedman, Emerson, Witzki, Howerter, & Wager, 2000). The visuospatial sketchpad is involved in verbal retention only to a negligible degree (see Baddeley, 2003) and will thus not be considered here.

The subsystem of working memory that is particularly associated with verbal short-term memory is the phonological loop. In this subsystem, verbal information can be stored for about two seconds in an abstract phonological code. Phonological

information refers to the sounds specific to a certain language, that is, to its phonemes. Phonemes are abstracted from the incoming speech signal and represent meaning discriminating sounds. For instance, the different possible phonetic realizations of an “r” (e.g., [ʀ] or [ʁ]) are all represented as a single phoneme /r/ in German, since the meaning of a word does not differ depending on how the /r/ is spoken. Whether an entity is a phoneme or not is language-specific. In German, the use of [s] or [θ] does not influence the meaning of a word, whereas in English it does (as in “sing” vs. “thing”). This means that the two sounds are separate phonemes in English but not in German. In German, a [θ] would just be a different realization of the phoneme /s/, as could be produced by someone who lisps. If one reads a word instead of hearing it, it is assumed that the graphemes are converted into phonemes (*grapheme-phoneme-conversion*, cf. e.g., Ellis & Young, 1988; Patterson, 1988). The phonological code assumed to underlie the phonological loop is abstract in the sense that both acoustic and visual verbal materials are encoded phonologically and stored in that format. Whether orthographic representations also play a role is not specified. In the language processing system, three different hierarchical levels are related to phonology: (1) the representation of single phonemes, which are connected to the (2) syllables (and the morphemes) in which they occur. These syllables are again connected to (3) the word forms. The word form level is the highest level that is assumed to be organized according to form information, not meaning. In Baddeley’s model, the phonological store is an entity that is separate from these representations but that relies on an equivalent code.

In order to retain information in the phonological store for longer than two seconds, this information has to be refreshed by subvocal rehearsal. Subvocal rehearsal can be conceived of as a kind of inner speech and is postulated to be part of the normal speech production or speech planning process. Suppose you are to keep in mind a list of things to buy at the butcher’s and you (silently) repeat it over and over until you have reached the shop and you can make your order. If the list is too long, that is, longer than what you can utter in two seconds, this will not work and you will lose some of the things you were supposed to buy.

In addition to prolonging retention intervals, the rehearsal mechanism serves one more function: it is used to recode visual verbal information into a format in

which it can enter the phonological short-term store; auditorily presented information enters the short-term store directly (see also Figure 1 below). Once information has entered the short-term store, it is stored in a modality-unspecific code. The encoded information does not differ for verbal materials that have entered the system as a string of sounds or of letters.

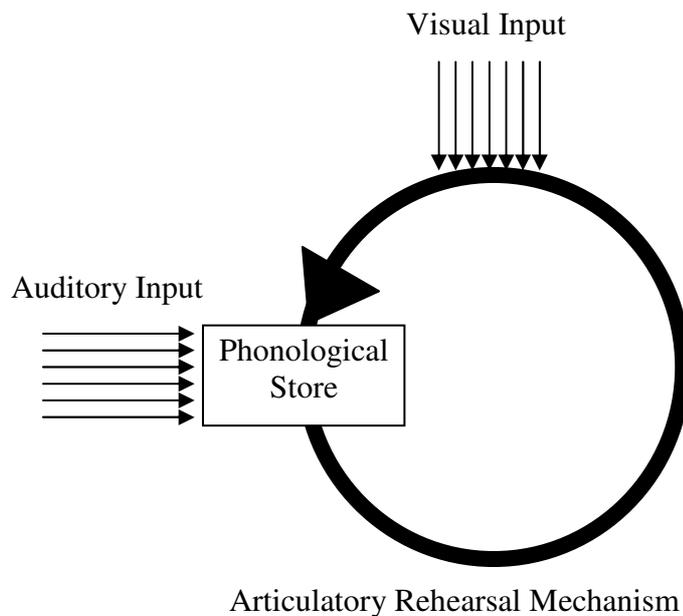


Figure 1: The phonological loop in Baddeley's (1986) model of working memory.

The paradigmatic task for the phonological loop is the serial recall of unrelated lists consisting of simple items (digits, letters, non-words, or words). The division into a passive store and an active rehearsal mechanism and the functions of these two make several predictions for such a task. In the following, I will briefly review some major findings from list recall and how they can be concluded from the architecture of the phonological loop. In general, one can serially recall about six items, as much as one can articulate in two seconds and thus as much as can be rehearsed and thereby be prevented from decaying.

Since rehearsal depends on subvocal articulation, the model predicts that recall performance depends on length or articulatory duration of the items. Indeed, it has been found that one can recall more short items than long ones (*word length effect*, Baddeley, Thomson, & Buchanan, 1975). In addition, phonological similarity between items (which is assumed to enhance interference and susceptibility to confu-

sion errors in the phonological store) decreases memory span (*phonological similarity effect*, e.g., Conrad & Hull, 1964), whereas semantic similarity does not impair performance equivalently (Baddeley, 1966; but see below for semantic influences on serial recall). Having extraneous speech sounds in the background also hampers serial list recall (*irrelevant speech effect*, Salamé & Baddeley, 1982). This is supposed to be due to interference that these sounds cause in the phonological store. If the rehearsal mechanism is engaged in concurrent articulation of irrelevant materials (such as “the the the the”), recall performance decreases substantially (*effect of articulatory suppression*, e.g., Murray, 1968). This is the case because the rehearsal mechanism is blocked.

Important evidence in favor of the phonological loop model comes from interactions between these different effects. Articulatory suppression blocks rehearsal and, as a consequence, recall performance with written materials is no longer influenced by phonological similarity (Murray, 1968). Phonological similarity is assumed to deteriorate retention in the short-term store. Yet as the rehearsal mechanism is needed to transpose written materials into a format that can enter the store (i.e., a phonological code), effects attributed to this store cannot influence recall if the items to be retained cannot enter it. With auditory presentation, articulatory suppression has no consequences on the phonological similarity effect (Murray, 1968). This is consistent with the assumption that auditory information enters the phonological short-term store directly, that is, without the involvement of the articulatory rehearsal mechanism. In addition, articulatory suppression eliminates the word length effect with both visual and auditory presentation. As indicated above, the word length effect is attributed to limitations of the rehearsal mechanism. However, these limitations do not play a role if the mechanism is blocked anyway (Baddeley et al., 1975).

With respect to the possible functions of the phonological loop beyond serial recall of lists, recent assumptions state that its contribution to language processing seems to be restricted to supporting vocabulary development (in both first and second language learning, Baddeley, Gathercole, & Papagno, 1998; Baddeley, 2003). Beyond language processing, the articulatory rehearsal mechanism as a kind of inner speech is assumed to serve as an internal self-cuing device when switching between tasks (cf. Emerson & Miyake, 2003).

Despite its generally good performance in accounting for effects in serial recall, there are some effects that cannot be explained in terms of the phonological loop and that suggest an influence of (linguistic) long-term memory representations. The phonological loop alone cannot account for findings that lexical-semantic factors influence performance in serial recall tasks. For instance, memory for word lists is better than memory for lists of pseudowords (the so-termed *lexicality effect*, Crowder, 1978). In addition, words of high frequency (i.e., words that one encounters rather often) can be better recalled than less frequent words (*frequency effect*, e.g., Watkins, 1977) and recall performance is better with concrete than with abstract words (*concreteness effect*, Walker & Hulme, 1999). Lists of words with high imageability lead to better recall performance than lists with low imageability words (*imageability effect*, Bourassa & Besner, 1994).

The explanation of these effects requires the inclusion of linguistic representations stored in long-term memory, such as entries in the mental lexicon or semantic features. In order to incorporate language representations and processes, recent versions of Baddeley's working memory model (since 2000, see Figure 2 below) include a new component: the *episodic buffer*.

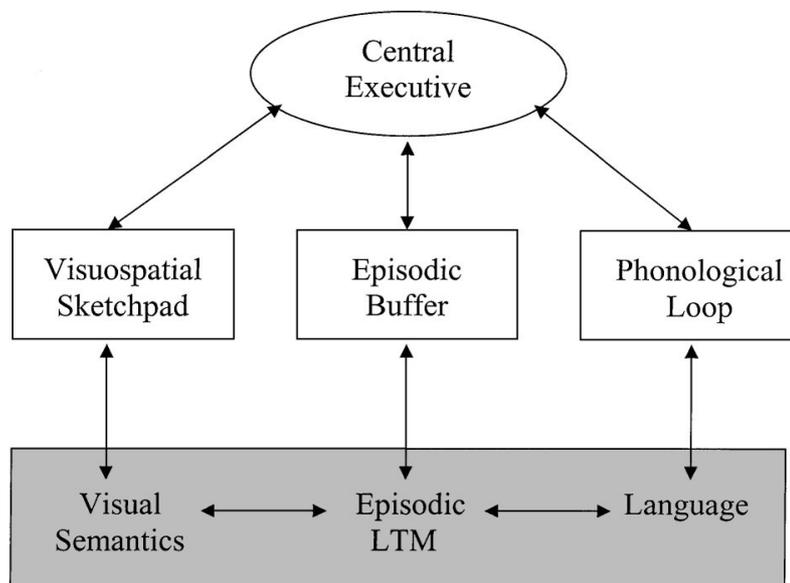


Figure 2: Baddeley's (2000) working memory model (white components) and its relation to long-term memory (grey component).

In this limited capacity system, information from long-term memory, such as linguistic information, is bound together with information stored in the phonological loop or

the visuospatial sketchpad and these different types of information can interact to form an episodic memory trace.

The addition of the episodic buffer brings about one more change. In the original version of the model the central executive was conceptualized as both a storage and an attentional control system. Now the episodic buffer essentially takes over the storage function of the central executive, which is thus re-conceptualized as an attentionally-based control system with no capacity for storage (cf. Baddeley & Logie, 1999). The buffer is assumed to play an important role in feeding information into and retrieving information from episodic long-term memory. It uses a common multi-dimensional code and provides a temporary interface between the slave systems and long-term memory.

The inclusion of the episodic buffer aims to account for interaction effects with information that is assumed to be stored in long-term memory. Although the episodic buffer provides an interface between long-term memory and the other subsystems of working memory, the model maintains the separation between these memory systems. Baddeley thus rejects the suggestion that working memory simply represents the activated portions of long-term memory (e.g., Cowan, 1995; 1999; see also Chapter 1.2 below) and the related view that the slave systems merely represent activations within the processes of visual and verbal perception and production (e.g., Jones, Macken, & Nicholls, 2004; Jones, Hughes, & Macken, 2006). The episodic buffer is rather conceived of as downloading information from long-term memory.

As indicated above, in the context of verbal short-term memory, the contribution of the episodic buffer is assumed to account for findings that linguistic factors such as lexicality or imageability influence immediate serial recall. Another problematic finding for the idea that serial recall is generally based on the phonological loop is that retention of syntactically structured materials (i.e., sentences) by far exceeds the capacity of the phonological loop. Sentences with a length of about 15 words or even more can be recalled. This is attributed to the contribution of the episodic buffer. As an attentionally limited temporary store, the episodic buffer formats links between unrelated concepts (chunks) and allows for simultaneous maintenance of these concepts. According to Jefferies, Lambon Ralph, and Baddeley (2004), sentence recall reflects contributions from automatic linguistic (syntactic and semantic) processes and attentionally limited working memory (i.e., the phonological loop and

the episodic buffer). However, how exactly the subsystems of working memory and the linguistic processes work together is not (yet) specified.

The relationship between short- and long-term memory is of growing importance in this line of research, as is shown by the addition of a new, more long-term memory related working memory component (i.e., the episodic buffer). Moreover, this relationship is of course of great relevance when considering possible interactions between verbal short-term memory and linguistic long-term memory information. In this context, two more types of models will briefly be introduced in the following subchapters: one prominent framework that conceives of working memory not as a separate subsystem but as the activated part of long-term memory (Cowan, 1995, 1999) and models that focus on the influence of linguistic long-term memory (Martin et al., 1999; Dell et al., 1997; Martin & Gupta, 2004).

1.2 Working memory as the activated part of long-term memory

Nelson Cowan's (1999) *Embedded-Processes Model of Working Memory* dispenses with a strict separation between long- and short-term memory (for a similar account cf. Oberauer, 2002). Instead, Cowan postulates that working memory is the currently activated part of long-term memory. He sees working memory as "cognitive processes that retain information in an unusually accessible state, suitable for carrying out any task with a mental component" (Cowan, 1999, p. 62). In the context of this study, I will outline solely the general framework of Cowan's theory. However, the model is more elaborate with respect to several aspects, for instance as to the capacity limit of the focus of attention (e.g., Cowan, 2001) or as to influences of attention or of sensory representations (e.g., Cowan, Saults, & Brown, 2004). With respect to the questions addressed here, the framework is of primary interest since the model on which the current study is based can be considered an elaboration of Cowan's account specialized for memory for linguistically structured materials.

According to Cowan, for all the information stored in long-term memory, there is a subset for which activation is heightened at a point of time (i.e., a *pre-activated part of long-term memory*). In addition, part of this subset is especially attended to. This information is in the *focus of attention* (see Figure 3 below).

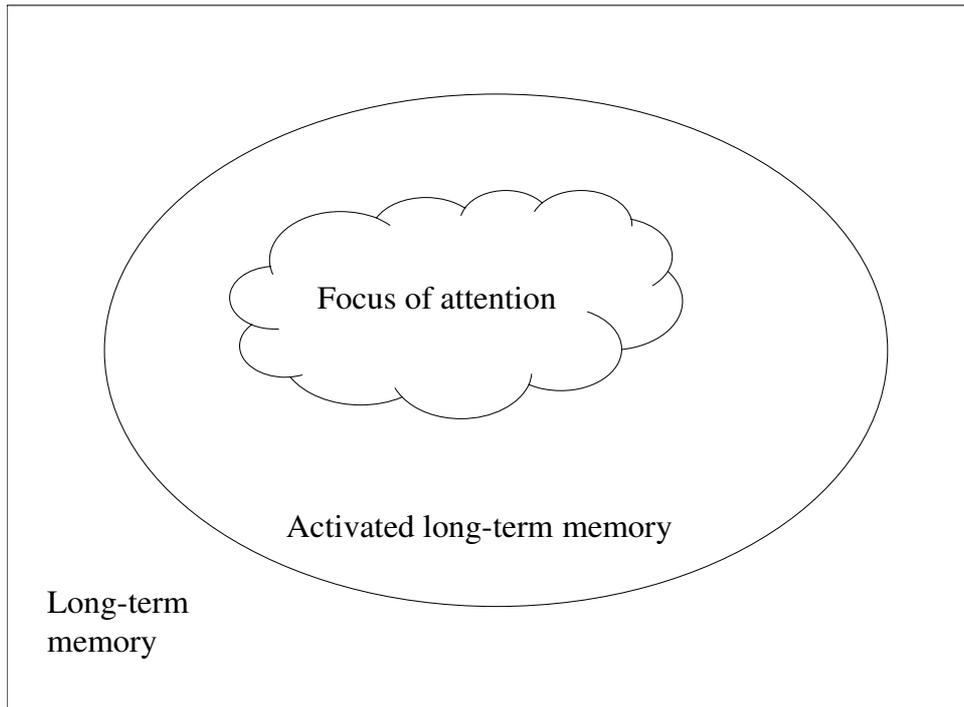


Figure 3: Cowan's (2001) tripartite memory model.

Cowan assumes that new connections are drawn between elements in the focus of attention and, possibly, in the pre-activated part of long-term memory. These links can be established such that they form new composites in long-term memory. In addition, “information that is not in an activated state can be thought of as in working memory [...], if there are cues in working memory that point to the item and raise the likelihood that it could be retrieved if necessary” (Cowan, 1999, p. 66). Working memory is not specifically attributed to one of the systems but is assumed to draw on information from all three memory components. The most accessible of these memory structures is the focus of attention.

This tripartite architecture of memory differs from previous accounts that have mostly separated short- and long-term memory (e.g., Atkinson & Shiffrin, 1968, or James', 1890, distinction between primary and secondary memory). Cowan's account implies instead that there is no separate working memory architecture: whether an item is in long-term memory or in the focus of attention depends simply on its state of activation and whether or not it is attended to. Thus, he assumes the same general principles of activation and de-activation to apply across all types of information, be they sensory or semantic. A mechanism that Cowan assumes to be shared among different kinds of codes is interference. Incoming stimuli that require a

particular kind of coding interfere with retained representations that use similar coding or similar neural mechanisms of activation. What kinds of codes can be differentiated is not specified. Cowan (1999) refers to the model proposed by Schneider and Detweiler (1987), who postulated multiple storage buffers (visual, auditory, speech, lexical, semantic, motor, mood, and context) and a central control module, but so far does not assume separate buffers himself.

According to Cowan, not all of the storage structures making up activated memory have capacity limitations. Instead, they are assumed to be limited because of memory decay and interference from subsequent stimuli. Only the focus of attention is assumed to have a fixed capacity limit in chunks. Four (plus/minus one) chunks can be stored in the focus of attention (e.g., Cowan, 2001). Cowan emphasizes that the critical thing about chunks in short-term memory is not their storage, since all information is already stored in long-term memory. What has to be maintained actually is the *relation* between elements. Normally, chunks are formed with the help of associations in long-term memory. The focus of attention might thus maintain pointers to the activated, stored information in long-term memory. By the time of recall, one could then shift one's focus to the respective stored information.

This account provides a framework for conceptualizing working memory in various fields. Among these are relations between visual attention, visual perception, and visual working memory. In the context of the present study, consequences for the relationship between language processing and verbal working memory are of special interest. The idea of a pre-activated part of long-term memory relates to priming mechanisms in language comprehension and language production. Models that differentiate this point further call themselves language-based models of short-term memory and are based on the assumption that short-term memory for verbal information draws largely on long-term knowledge representations of the language processing system. These models will be briefly introduced in the next section.

1.3 Language-based models of short-term memory

Similar to the model sketched above, language-based models of short-term memory share the assumption that there is no strict separation between long-term memory and short-term memory. Roughly speaking, these models can be seen as an elaboration for a certain part of long-term knowledge: language. According to these models,

short-term memory for verbal materials builds on the same long-term representations as normal language processing: “phonological, lexical, and semantic representations are activated in both language processing as well as in verbal short-term memory.” (Martin et al., 1999, p. 5).

Two related views shall be introduced here. Both accounts share the assumption that verbal short-term memory depends on word perception and word production but they differ in how they have implemented the retention components. In the model proposed by Martin et al. (1999) the retention components are instantiated by buffers for phonological and lexical-semantic information which are each connected to the respective levels of the language processing system (see Figure 4 below).

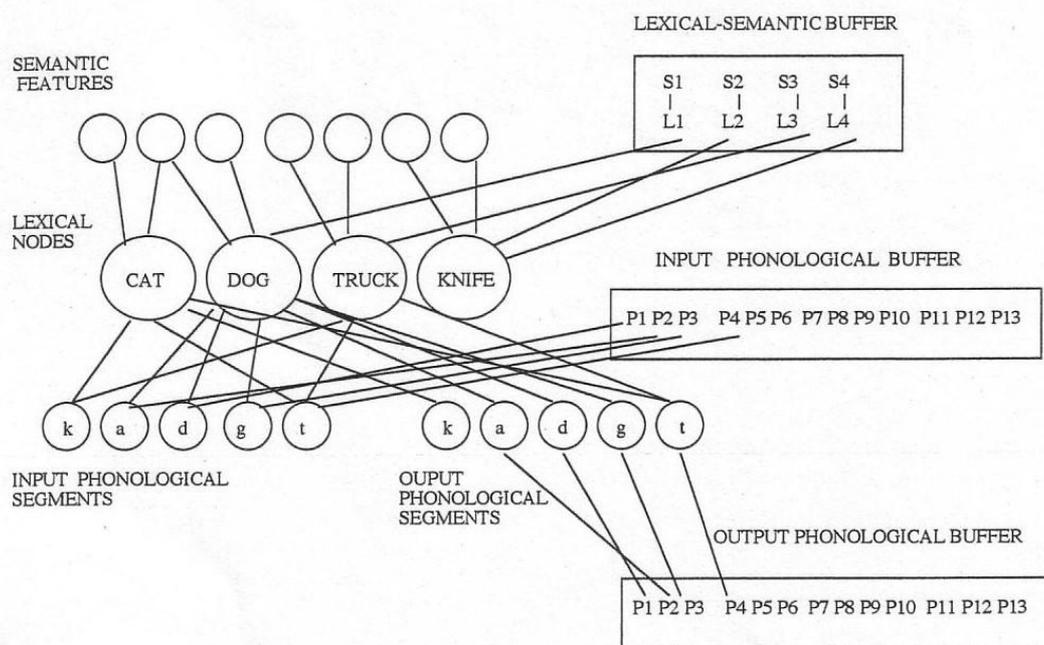


Figure 4: Martin et al.'s (1999) conception of verbal short-term memory. The left side shows the language processing system, the right side shows the respective short-term buffers.

The organization of the model into buffers associated with different levels of the language processing system goes back to Barnard (1985). The model by Martin et al. (1999) applies Gary Dell's interactive activation model of word processing (originally conceived as a language production model, e.g., Dell, 1986; Dell & O'Seaghda, 1992; the left side of Figure 4), in which different representational levels are assumed. Word perception and production share lexical and semantic nodes but have different representations for input and output phonology. Between these levels excitatory bidirectional links exist, that is, activation spreads from semantic to lexical

nodes and vice versa and from lexical to phonological nodes and back (for a different account of word production see e.g., Levelt, 1989; Levelt et al., 1999).

If a word is to be produced, the semantic features that are associated with the intended message activate the lexical nodes that are connected with them. Thereby not only one lexical entry is activated but also related ones. For instance, when intending to name a lion, it is probable that tiger also gets some activation since lion and tiger share semantic features. Each lexical node that gets activation feeds activation forward to the phonological level so that the phonemes it consists of get activated. From the lexical nodes activation is also sent back to the semantic level. Thereby the relevant semantic features get reinforced. The phonological nodes feed back activation to the lexical level, too. Feed back increases the activation differences between the intended word and related ones and thereby increases the probability that the right word (phoneme, etc.) is selected.

The architecture of Dell's model is called "globally modular and locally interactive", indicating its organization in representational levels and the interactive activation between levels. Other models of word production do not agree that there are feed back connections between the phonological and the lexical network (e.g., Levelt, 1989; Levelt, Meyer, & Roelofs, 1999).

To include a short-term retention component into the model, buffers have been added that are connected to the respective nodes on the lexical-semantic and the phonological levels. Activation spreads from the representations of the language perception and production system to the buffers and vice versa. In contrast to the nodes in the language production network, the buffers themselves are assumed not to interact with each other. Nevertheless they reflect interaction processes between lexical-semantic and phonological information in the long-term knowledge store, that is, in the language processing system. In this account, word retention and word processing build on the same representations.

The principal implication of such an assumption is that factors that influence word perception or word production should also influence word retention. The model thus predicts that both phonological and lexical-semantic representations are involved in short-term memory tasks. Such an involvement has indeed been observed. Experiments with unimpaired subjects and neuropsychological patient studies can be cited as evidence for shared representations in word processing and word retention.

Some of these studies will be summarized in the next paragraphs (see also Chapter 1.1).

It has long been demonstrated that phonological factors such as word length or phonological similarity influence short-term serial recall (e.g., Baddeley et al., 1975; Conrad, 1964). If items that are supposed to be recalled are phonologically similar, recall performance is worse. For lexical features an involvement in serial recall tasks can also be found. Memory span is greater for words than for non-words (lexicality effect, Crowder, 1978). This reflects the circumstance that non-words can only be represented as a novel string of phonemes but words get additional activation from an entry in the mental lexicon and can be stored in a lexical format. In addition, one can recall more high frequency words than words with low frequency (frequency effect, e.g., Watkins, 1977). In Dell's language processing model, word frequency affects the activation levels in the lexical network. High frequency words are per se activated more highly than low frequency words. Thus, less activation is needed for these words to be selected and they can send more activation to their associated representation in the lexical-semantic short-term store. These characteristics account for a recall advantage for high frequency over low frequency words.

In addition to phonological information and the mental lexicon, semantic features influence recall performance. More high imageability than low imageability words can be serially recalled (imageability effect, Bourassa & Besner, 1994). Imageability refers to the ease with which a word produces a mental image irrespective of whether it is concrete or abstract. However, there is great correspondence between the high versus low imageability and the concrete versus abstract distinction. In line with this overlap, concrete words can also be serially recalled more accurately than abstract words (concreteness effect, Walker & Hulme, 1999; Allen & Hulme, 2006).

Both imageability and concreteness effects reflect semantic influences. Concrete words and high imageability words are assumed to have a richer semantic representation or a semantic representation that depends less on context than low imageability words (cf. e.g., Saffran, Bogyo, Schwartz, & Marin, 1980). In the current model this would imply that more semantic feature nodes are connected with the lexical nodes of both highly imageable words and concrete words than with the lexical nodes of low imageable words and abstract words (Dell, Schwartz, Martin, Saffran, & Gagnon, 1996; for a related implementation in a parallel distributed process-

ing approach see Plaut & Shallice, 1993). Therefore, the lexical entry of a high imageability word gets activation from more nodes on the semantic level than a low imageability word and in consequence, the representation in the lexical-semantic short-term buffer also gets more activation. In addition, retrieval of the lexical entry of a high imageability word is less vulnerable. Allen and Hulme (2006) attribute the recall advantage of concrete over abstract words to the ease of eliciting an appropriate word form on the basis of the semantic representation.

Neuropsychological patient data provide evidence for language-based accounts of short-term memory in two respects. A first approach analyzes performance of patients with selective neuropsychological impairments in order to provide evidence for the assumption that the contribution of phonological and lexical-semantic codes can be dissociated not only in language processing but also in short-term memory. For instance, patient EA, who suffers from a selective deficit in phonological short-term memory, demonstrated a normal lexicality effect in her memory span but did not show effects of phonological similarity or word length (e.g., Martin, 1993). This suggests impaired phonological retention in the presence of unimpaired lexical processing in the context of short-term memory. Complementary patterns are found in patients AB and ML, who demonstrated preserved phonological similarity and word length effects but no lexicality effect (Martin & He, 2004; Martin, Shelton, & Yaffee, 1994).

The second approach addresses the issue of comparable impairments in processing and retention of verbal materials. If short-term memory for verbal materials depends as heavily on language processing as language-based short-term memory models suggest, patients who have impairments in language processing should show comparable impairments in related short-term memory tasks. Martin et al. (1999) observed such associated deficits in naming and word repetition in patient MS. His performance showed not only that his naming problems resulted in retention deficits, but also that his error patterns were comparable in both tasks. He often compensated his naming deficits with semantically adequate descriptions and produced phonological approximations. These strategies could also be observed in his list recall performance. For instance, MS recalled the list 'lobster, castle, bagpipe' the following way: "losser – the thing you eat, the place the kings go in, it comes from the place where men wear the same thing as women" (Martin et al., 1999, p. 19).

Similar to the approach of Martin et al. (1999) is an account proposed by Dell et al. (1997; see also Martin & Saffran, 1997; Martin & Gupta, 2004), which is based on the same language production model. These authors also assume that normal language representations are involved in verbal short-term memory tasks. Although the model relies on the same basic model of language production (Dell & O'Seaghda, 1992; see also the left side of Figure 4 above) it differs from Martin et al. (1999) in that no separate buffers are assumed in order to model short-term memory processes. Retention of verbal information is represented by *prolonged activation* of the respective information in the language processing system instead. This is similar to the general idea of Cowan's (1995, 1999) model (see also Chapter 1.2 above) and can be viewed as a specification of Cowan's account with respect to language.

The evidence for language-based accounts of short-term memory cited above also applies to Dell et al.'s (1997) theory. Since the two models are so closely related, there is no clear evidence so far whether it is more plausible to assume short-term storage as prolonged activation or in terms of associated buffers. Despite their slight differences both models largely head in the same direction and promote a language-based view of short-term memory.

As can be seen from this short review, there is a good empirical basis for the assumption that representations are shared between word perception, word production, and word retention. However, the accounts described so far focus on single word processing. Parallels between language processing and short-term memory should just as well be observable with more complex stimuli such as sentences. Sentences are structured conceptually and syntactically and their processing involves thus even more long-term knowledge representations than processing of single words. Therefore, the study of short-term memory for sentences should be even more challenging for theories that do not or minimally incorporate linguistic long-term knowledge. The following subchapters will focus on theories of short-term memory for sentences.

1.4 Short-term memory for sentences

Memory for sentences differs from memory for words (in isolation or in unrelated lists) in several respects. The most obvious difference concerns the number of unrelated words and the number of words in a sentence that can successfully be recalled.

While in a list about six words can be repeated correctly (in serial order), sentences of about 15 words or more can be recalled without substantial problems.

Another obvious difference concerns the relation between the words on a list and the relation between the words in a sentence. Normally the words on a list are unrelated, whereas the words in a sentence are structured syntactically and conceptually. In addition, while word lists are usually constructed such that all items come from the same word category (predominantly nouns), a sentence consists of words from various categories: nouns, verbs, adjectives, and function words such as determiners, connectives, and so on. Since any sentence needs a verb in order to be grammatical, omission errors for verbs occur very rarely in sentence recall. Moreover, word order in a sentence is a priori very restricted (particularly in English). Thus, in a sentence like “The little boy loved his cute cat.” there are hardly any options for order errors to occur, if the result is supposed to be a grammatical sentence. Obviously, such restrictions do not apply to a word list, such as, “theme, litter, boy, love, hiss, cue, cat”. These structural differences suggest that retention of syntactically structured materials might engage mechanisms that differ from mechanisms that subserve serial list recall. A comprehensive theory of verbal short-term memory needs to account not only for performance in serial recall tasks but also – and in particular – for sentence recall performance.

It seems plausible that the structural differences between sentence and list *recall* reflect structural differences between sentence and list *processing*. An impressive example for differences in processing of lists and sentences comes from a neuropsychologically impaired patient, Mr Clermont (Nespoulos, Dordain, Perron, Bub, Caplan, Mehler, & Lecours, 1988). Mr Clermont had no problems producing, reading, or repeating words in isolation or in unrelated lists but produced syntactic and morphosyntactic errors in sentence production (predominantly omission and erroneous production of function words). Astonishingly, Mr Clermont read the words of a sentence appropriately when they were written vertically one word below the other (i.e., listlike) but only until he realized that he was actually reading a sentence. From that point on he read the list like a sentence and produced the same error patterns that he normally produced when reading sentences.

If the differences between list and sentence retention can indeed be attributed to characteristics of language processing, then language-based theories of short-term

memory in particular need to account for sentence recall characteristics. In the following subchapter, theories will be reviewed that focus on sentence recall. First of all, early theories of sentence memory focus on the distinction between memory for surface information and memory for the gist of a sentence. After these theories, the conceptual regeneration hypothesis will be introduced, which conceives of sentence memory as a regeneration process based on a conceptual representation of a sentence's message. The final theory presented here proposes that short-term memory for sentences follows from the necessities of sentence processing. This theoretical account inspired the present study.

1.4.1 Memory for surface and gist information as a function of a sentence's state of processing

Traditionally, memory for sentences has been divided into a short-lived verbatim surface representation and a long-lasting gist representation. The idea is that the most recently processed sentence is still available in a surface format that is assumed to be similar to the ordered short-term representation of a word list. In contrast, what stays in memory from previous parts of discourse is conceptual information only.

Findings from recognition and recall studies in the late 1960s and the 1970s suggested that clauses and sentences were the critical processing units that determined shifts in memory representations. For instance, Caplan (1972) presented his subjects with complex (two-clause) sentences and afterwards with a probe word. They had to decide whether the probe word had been part of the sentence or not. He observed that recognition latencies were shorter for probe words from the second clause than for probe words from the first clause. This was the case even though the target word in the sentence and the following words were the same in both conditions. The only difference was whether there was a clause boundary or not. Examples (2a) and (2b) illustrate this manipulation (the target word is italicized):

(2a) Now that artists are working in *oil*, prints are rare.

(2b) Now that artists are working fewer hours, *oil* prints are rare.

Caplan concluded that surface representations such as lexical items from the most recent clause were more accessible than from prior clauses. In an earlier study, Sachs

(1967) presented her participants texts that each contained a target sentence. After text presentation, a probe sentence was presented for which participants had to decide whether it had been in the text or not. If not, they had to indicate in what way it differed from a sentence that was indeed part of the text. The probe sentence could bear different relationships to the target sentence in the text: it was exactly the same one, was changed in content, was changed in its syntactic structure (active vs. passive voice), or word order had been changed but without resulting in a change in content (termed formal changes). Sachs (1967) observed that after a brief delay (80 interpolated syllables), surface changes were detected less accurately than changes in content (see also Figure 5 below).

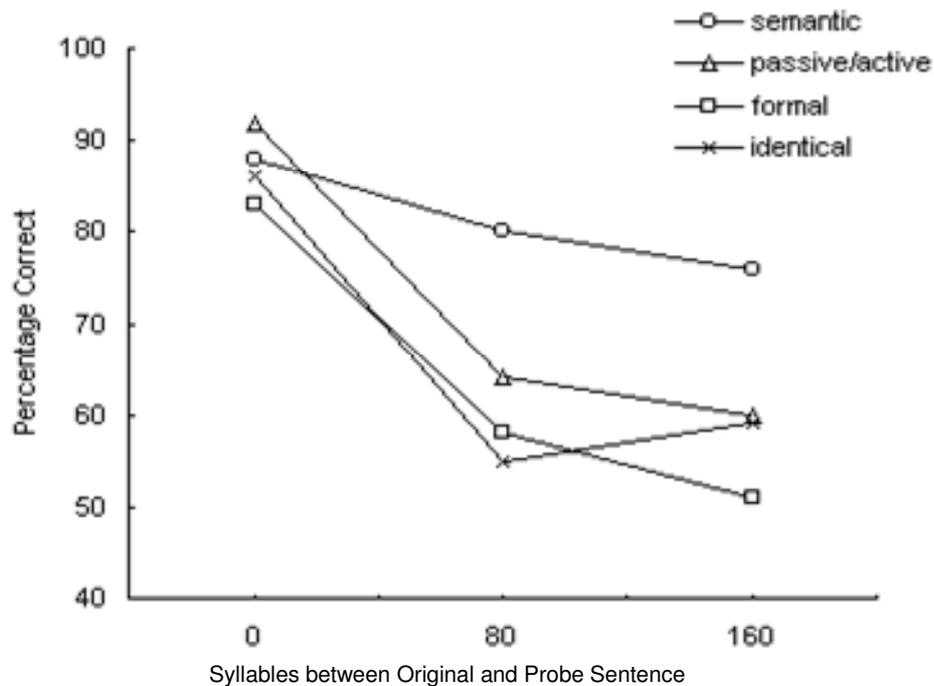


Figure 5: Percentage of correct judgments (“identical” or “changed”) for each probe type as a function of the number of syllables between original and probe sentence in Sachs’ (1967) experiment.

In a subsequent study, Sachs (1974) compared auditory and visual presentation of the stimuli and tested shorter delays between the original sentence in the text and the probe sentence (0, 20, 40, and 80 syllables between original and probe sentence as compared to 0, 80, and 160 syllables in the previous experiment). In addition, she

included a further type of probe sentence: lexical changes, that is, words in the original sentences had been replaced with synonyms. Her results are depicted in Figure 6 below.

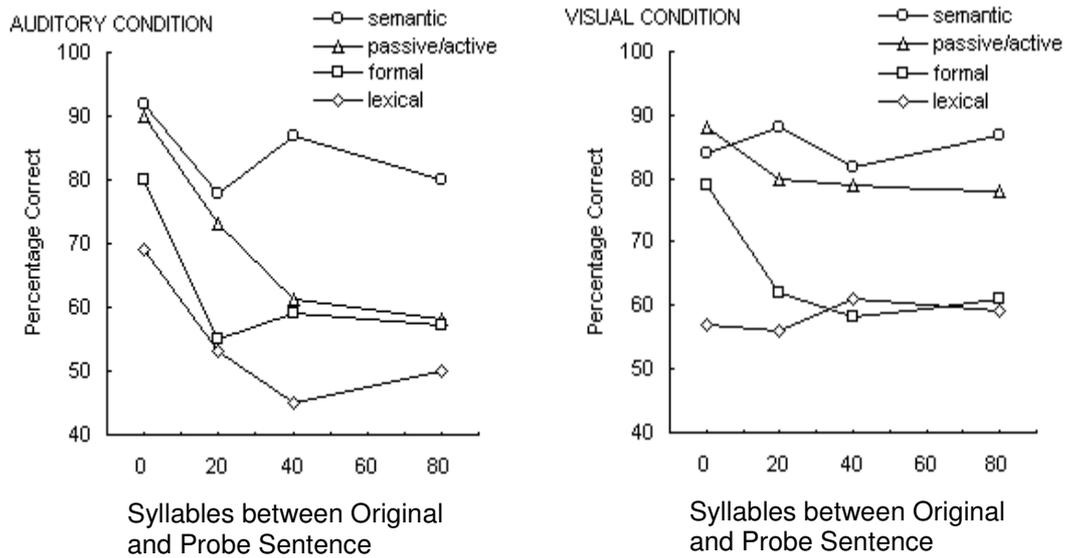


Figure 6: Accuracy in judging changes for each probe type as a function of presentation modality and of the number of syllables between original and probe sentence (Sachs, 1974).

In this experiment, Sachs' earlier findings were modulated such that syntactic changes were detected more reliably than others. With auditory presentation, this difference was observable after a brief delay (20 syllables). With visual presentation, this difference was observable even after 80 syllables had been read between the original and the probe sentence. In addition to modality-specific differences, Sachs' (1974) results indicate that syntactic information was available longer than other types of surface information.

Similar patterns of results were obtained in recall studies. Jarvella (1971, 1972) found that surface representations of a sentence were no longer available (or only rudimentary so) as soon as the sentence had been entirely processed. He had his participants listen to texts. Whenever presentation of these texts was interrupted, participants were instructed to write down (Jarvella, 1971) or to recall orally (Jarvella, 1972) as much of the previously perceived discourse as possible. Jarvella's critical manipulation concerned the final three clauses just before the interruption. In these structures, the middle clause could either be connected to the previous or to the final

clause. As a consequence, in the former case it was no longer part of the most recent sentence, whereas in the latter case it was. Middle clauses with exactly the same wording were compared in both conditions (see example (3a & 3b) below, Jarvella, 1971, p. 411).

(3a) Taylor did not reach a decision until *after he had returned to Manhattan*.
He explained the offer to his wife.

(3b) With this possibility, Taylor left the capital. *After he had returned to Manhattan*, he explained the offer to his wife.

An analysis of recall performance for the three clauses revealed that the final clause was always recalled more accurately than the previous ones and, critically, that recall performance for the middle clause strongly depended on its state of processing. That is, whereas performance for the middle clause was as low as for the first clause when they belonged to the same sentence, recall for the middle clause improved significantly when it was connected to the last one and thus part of the most recently processed sentence. Figure 7 depicts Jarvella's (1972) results.

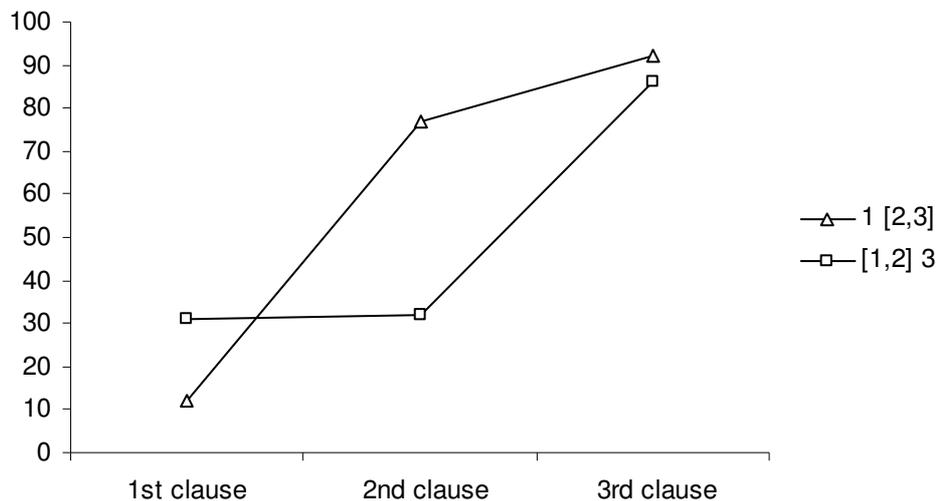


Figure 7: Proportion of correctly recalled words per clause as a function of syntactic structure in Jarvella (1972). “[1,2] 3” refers to the condition in which the first and the second clause form a sentence together and “1 [2,3]” refers to the condition in which the second clause forms a sentence with the third (i.e., the final) clause.

This general pattern of data has been replicated under various conditions (for an overview cf. Jarvella, 1979), for instance with both auditorily and visually presented texts or with a general recall instruction and with cued recall (the first word of the first of the three critical clauses being the recall cue). Since Jarvella kept the wording of the middle clauses constant, the observed differences can be attributed to the differences in the state of processing. From his studies, Jarvella concluded that the surface form of a sentence is only maintained as long as it is actively processed. Otherwise only the gist of the sentence is represented. Critical stages are clause boundaries and, even more so, sentence boundaries. It is worth noting that Jarvella (1971, 1972, 1979) did not distinguish different types of surface information. In contrast to Sachs' (1967, 1974) experiments, no conclusions can be drawn concerning potentially different decay rates of lexical or syntactic information.

A different account of verbatim sentence recall was proposed in the 1990s by Mary Potter and Linda Lombardi (1990, 1998; Lombardi & Potter, 1992). Based on results obtained by von Eckardt and Potter (1985), they rejected a decisive role for surface information in any form of sentence recall. Similar to Jarvella (1971, 1979) and Caplan (1972), von Eckardt and Potter (1985) found that probe recognition performance deteriorated at clause boundaries – but both with probe words and probe *pictures*. After reading a two-clause sentence, their subjects were presented either with a word or with a picture of a concept mentioned in the sentence. Probe recognition latencies for both pictures and words were shorter, if the respective concept belonged to the most recent clause than when it belonged to the previous one. It is hard to argue that the clause recency effect with *pictorial* probes goes back to a more exact representation of *linguistic* surface information. Von Eckardt and Potter (1985) attributed the performance difference to better availability of the *conceptual* information of the most recent (part of the) sentence. Potter and Lombardi (1990) explored the view expressed by von Eckardt and Potter (1985) and argued, in contrast to Jarvella (1971), that conceptual information was the basis of both verbatim and gist memory for sentences and thus tied short-term sentence recall more closely to verbal long-term memory. Their influential *conceptual regeneration hypothesis* will be described in more detail in the following subchapter.

1.4.2 *The conceptual regeneration hypothesis*

Potter and Lombardi (1990, 1998; Lombardi & Potter, 1992) proposed a language-based view of sentence memory. They noted that not only gist memory but also verbatim short-term memory for sentences is based on a propositional-conceptual structure which is generated during comprehension of the respective sentence. Starting from this representation, the sentence is *regenerated* via mechanisms involved in normal sentence production (instead of retaining the sentence as a surface string of words).

As there is often more than one way to express a propositional structure (e.g., with near-synonymous words such as “leap” or “jump” or in active or passive voice), a propositional-conceptual representation alone is not sufficient to guarantee verbatim recall. In addition, Potter and Lombardi (1998) assume two priming mechanisms to add to the regeneration process. First, the syntactic structure is primed while processing the sentence and, second, the activation of entries in the mental lexicon that have been part of the presented sentence is higher than activation of semantically related – but not presented – competitors. Therefore, these items have a higher probability of being selected for regeneration. Taken together, a propositional representation and syntactic and lexical priming provide a sufficient basis even for verbatim sentence recall without the need to assume a surface short-term store independent of language representations.

Potter and Lombardi (1990) empirically supported their account of sentence memory via a paradigm that combines short-term recall of sentences and lexical priming. Since this intrusion paradigm has been applied in the present study, I will describe it in more detail here. In Potter and Lombardi’s (1990) experiments, participants had to recall sentences that contained a *target word* which was not an optimal fit for the sentence content, for example:

- (4) The knight rode around the *palace* searching for a place to enter.

As distractor task, a word list which had been presented either before or after the sentence had to be retained. Participants had to decide whether or not a subsequently presented probe word had been on the list. In half of the trials, one item on the list was replaced by a *lure word* (in the above example: *castle*). This lure word was se-

manically related to the target word and, in addition, was a better fit for the sentence content or was more strongly associated with the situation described. This manipulation often led to erroneous recall: even in those cases in which the lure word was not presented subjects sometimes produced *spontaneous intrusions*, and the proportion of intrusions increased significantly if the lure word was actually presented as part of the list (*induced intrusions*). Potter and Lombardi (1990) attributed the occurrence of spontaneous intrusions to the influence of the sentence context (i.e., the propositional-conceptual structure). They assumed that sentence context suggests the better fitting word for recall and attributed the induced intrusions to the influence of the sentence context *and* to higher activations of primed items in the mental lexicon.

With respect to syntactic information, Lombardi and Potter (1992; see also Potter & Lombardi, 1998) suggested that it is involved in sentence recall only implicitly via priming mechanisms. Lombardi and Potter (1992) demonstrated that morphosyntactic features of the verb did not constrain intrusions from verbs with different morphosyntactic features. Instead of nouns they used verbs as target and lure words and varied whether the lure verb was compatible with the morphosyntactic (subcategorization) characteristics of the target verb in the sentence. That is, they used verbs that required different complements. The target word was compatible with two different structures, for instance with a double object structure and with a prepositional phrase structure, while the lure verb was compatible with only one of these structures (see example (5) below).

(5a) The rich widow is going to *give* a million dollars to the university.

(5b) The rich widow is going to *give* the university a million dollars.

(5a with intrusion) The rich widow is going to *donate* a million dollars to the university.

(5b with intrusion) *The rich widow is going to *donate* the university a million dollars.

If surface syntactic structure was directly represented in short-term memory, the lure verb should not substitute the target verb in the incompatible sentence. Yet this was not the case. Intrusions occurred and participants adapted the syntactic structure of the sentence to the morphosyntactic constraints of the verb they used. However, there

was still a tendency to reproduce the sentence with the syntactic structure of the presented sentence. Lombardi and Potter (1992) attributed this to syntactic priming mechanisms. In a subsequent study, Potter and Lombardi (1998) provided further evidence for their assumption that syntactic priming underlies the structural similarity of perceived and reproduced sentence: syntactic surface structure of a first sentence primed recall of a subsequently presented sentence (see also Bock, 1986).

Altogether, Potter and Lombardi state in their conceptual regeneration hypothesis that memory for sentences is driven by a propositional-conceptual representation, with lexical and syntactic priming contributing in addition, while phonological or morphosyntactic short-term representations do not seem to play a decisive role.

1.4.3 Memory for sentences as a function of sentence processing

Rummer and Engelkamp (2001, 2003) modified the conceptual regeneration hypothesis. In line with Potter and Lombardi (1990) they assumed that conceptual representations built in the course of sentence comprehension are fundamental for subsequent sentence recall. Yet they questioned the dismissal of the contribution of so-called surface information and assumed (and provided evidence for) an additional influence of phonological information on sentence recall.

Based on these preliminary results, Rummer and Engelkamp (in press, see also Rummer, 2003) have proposed a model of short-term memory for verbal materials focusing on (but not restricted to) sentences. This model aims to go beyond both the word-based short-term memory models (Martin et al., 1999; Dell et al., 1997) and the conceptual regeneration hypothesis (Potter & Lombardi, 1990; 1998). With these models Rummer and Engelkamp (in press) share the assumption that language processing and verbal short-term memory are strongly intertwined.

One basic idea of Rummer and Engelkamp's (in press) model is that those representations that are activated (or generated) in the course of comprehending a sentence are maintained for some time (for a specification of "some time" with respect to different types of representations see below) and can be preferentially accessed in sentence regeneration. For this purpose, the model combines a (simplified) model of sentence comprehension with a (simplified) model of sentence production and assumes that on each level the input representations are connected with the output representations (see Figure 8 below).

To understand a sentence like (6), one has to analyze the phonological structure of each word, access the right entries in the mental lexicon (in example (6) e.g. “cost” instead of “coast”), access their morphosyntactic features (e.g., word category: is “costs” in (6) a noun or a verb?), and their meanings (e.g., select among the different meanings of “union”). Furthermore, one needs to find out which words belong together syntactically (i.e., form a phrase, e.g. the noun phrase “the trade union leader”), extract the overall syntactic structure (did the trade unionist criticize the efforts *in order to* reduce labor costs or did he criticize the efforts *that aimed to* reduce labor costs?), assign thematic roles (who did what to whom?), and finally make use of all of this information to get a representation of what the whole sentence means. These operations result in potentially stored (by-) products on all representational levels (see Figure 8 above).

Since sentences are normally rapidly understood, these processes need to operate very fast. Theories of sentence comprehension agree that comprehension is incremental, that is, sentence interpretation is not delayed until a clause or a sentence boundary is reached but occurs well before. Yet these theories differ in the degree to which they assume language comprehension to be parallel and interactive or modular (for an overview of different theories of sentence comprehension cf. e.g., Harley, 2001; Hemforth & Konieczny, 2002).

Some theories suggest that at first the syntactic structure is analyzed, independent of the sentential context and the meanings of the individual words. Only in a second stage, an interpretation is built on the basis of the latter types of information. The two interpretations converge – in which case the sentence is understood – or they differ – in which case the sentence has to be reanalyzed (e.g., Ferreira & Clifton, 1986; Fodor, 1983; Frazier & Rayner, 1982; Frazier & Clifton, 1996).

According to other approaches, contextual, lexical, or structural frequency constrains sentence processing right from the beginning. This implies, for instance, that an implausible interpretation is less likely to be made than a plausible one (e.g., Altmann, 1998; Altmann & Steedman, 1988; Gibson & Pearlmutter, 1998; MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell, Tanenhaus, & Garnsey, 1994). So far, the current sentence recall model does not (need to) distinguish between these different approaches.

As is assumed by the conceptual regeneration hypothesis, the message of the sentence, which is the end point of sentence comprehension, is the starting point of sentence regeneration. In contrast to Potter and Lombardi (1990), Rummer and Engelkamp's (in press) model states that representations on *all* levels that have been pre-activated during sentence comprehension (phonological, lexical-semantic, etc.) contribute to sentence recall. Due to pre-activation, these representations have a higher activation level and are therefore most probable to be selected when the message has to be put in words (or structures) again. The idea that representations used in comprehension influence the use of representations in production is also part of the *interactive alignment* approach put forward by Pickering and Garrod (2004, see also Garrod & Pickering, 2004). Their model principally addresses the processes involved in dialogues. It states that interlocutors reciprocally activate each other's linguistic representations while listening and talking. Yet the assumption of connections between comprehension and production on each representational level (phonological, semantic, etc.) also applies to intra-personal representations. To illustrate the right-hand side of Figure 8 above, I will give a sketch of processes involved in sentence production in the following paragraphs.

An influential model of language production has been proposed by Levelt (1989). Although it is not without controversy (for different accounts see e.g., Dell & O'Seaghdha, 1992; Herrmann & Grabowski, 1994; Caramazza & Miozzo, 1997; Schade, 1999, and the model described in Chapter 1.3), I will describe the process of language production on the basis of this model. According to Levelt (1989), normal sentence production starts with generating a (preverbal) message (which in the case of sentence *reproduction* already exists) that then has to be formulated.

The first step of formulation is grammatical encoding (accessing lexical entries and building syntactic structure), which results in a surface structure of the message. For word selection, Levelt assumes two steps (while others assume only one, e.g., Caramazza & Miozzo, 1997). First, concepts activate an element that includes (morpho-) syntactic information, the *lemma* (Levelt et al., 1999). The lemma is specified with respect to word category, gender, and subcategorization information (e.g., whether a verb is transitive or intransitive). Not only the target concept but also semantically related concepts are activated so that these, too, send activation to their associated lemmas.

Levelt et al. (1999) assume that only the most highly activated lemma, that is, the selected lemma can send activation to the next processing level, the word form level. Other models of language production postulate that any lemma that is activated can also send activation to its associated word form. The respective lexical entry activates the phonemes it consists of. The result of this phonological encoding process is an articulatory plan (or possibly a writing motor plan). With the exception of activation flow from concepts to lemmas, Levelt et al. (1999) assume each component to work autonomously and activation flow to be restricted to already selected entries (i.e., only the ‘winner’ activates its related low level elements, the competitors do not).

This position is not generally accepted; other accounts assume language production to be more interactive (e.g., Dell, 1986; Dell & O’Seaghdha, 1992; Schade, 1999). However, for the present state of the short-term memory model it is not important whether one or more word forms can be activated or whether activation flows only top down or, in addition, bottom up via feed back connections. Thus, I will not discuss the empirical evidence for or against either of these positions.

A monitoring process is assumed to watch over speech production. Before articulation, the speech plan undergoes covert rehearsal and is checked by the monitoring component. If an error is detected, the monitor initiates a repair. In addition to covert rehearsal, the monitor controls overt speech and, in case of an error, initiates overt repairs. Nevertheless, the monitor does not detect all errors. That is lucky for psycholinguists, since speech errors have often been taken as indicators for the existence of different levels of representation.

When producing an utterance that is meant to convey the message of sentence (6) above, errors could arise on each step. For instance, a related word could be selected that does not exactly match the intended message (e.g., “condemns” instead of “criticizes”), the functional elements could be uttered in the wrong order (e.g., “The trade union leader criticizes *labor costs* to reduce *the efforts*.”), a morphosyntactic agreement error could occur (e.g., erroneous number agreement between noun phrase and verb: “The trade union leader *criticize*...), or the speaker might not be able to retrieve the correct phonological form and thus produce a phonologically similar word (“The *trait* union leader...”).

Many theories of sentence production investigate mechanisms of syntactic structuring or lexical selection. An important finding in this respect is that speakers tend to reuse syntactic structures they have produced or perceived before. Such syntactic priming can be observed even though there is no lexical overlap between prime and target (e.g., Bock, 1989) but is stronger when the same verb is used (Pickering & Branigan, 1998). When lexical entries need to be retrieved, it is assumed that several items compete for selection. Normally, the lexical entry that maps best onto the intended message is selected but noise (i.e., random activation) in the production system can sometimes lead to the retrieval of a related entry (for an overview of studies on lexical selection in language perception and production cf. Zwitserlood & Bólte, 2002).

In contrast to sentence production, the input in sentence reproduction does not only come from the message level. Potter and Lombardi (1990, 1998) suggested that lexical entries that have been activated during sentence comprehension are still activated more highly than their competitors and that due to syntactic priming the syntactic structure of the input sentence is most probable to be reused (see 1.4.2 above). Rummer and Engelkamp (in press) extend this to the assumption that *any* relevant information activated during sentence comprehension can potentially be used for sentence regeneration. Whether a particular type of information (e.g., phonological or semantic information) can indeed contribute to a certain recall task, depends on the time course of activation and decay, which is assumed to differ between the types of information included in the model.

As an elaboration of Rummer and Engelkamp's model, Rummer and Schweppe (in prep.) have explicitly attributed such time course differences to characteristics of the language processing system and to the role that each information type plays in language comprehension and production. In other words, critical for the duration of activation is the 'sustainability' of the information. The longer a certain type of information is needed for understanding further parts of discourse or for producing subsequent parts of discourse, the longer this information will be retained in short-term memory. That is, in addition to message-level and lexical-semantic information, also phonological (for a short period of time) and morphosyntactic information (for a longer period of time) should be maintained. The general idea behind this assumption is that the characteristics and requirements of the language processing

system determine the contents of verbal short-term memory and the capabilities of the stored representations, in particular their decay functions. Language processing requirements determine what kind of representations will be retained and for how long they will be maintained. For instance, phonological information is necessary for initial sentence processing (feeding lexical access) but not for understanding subsequent sentences, since then access to lexical representations should be sufficient. Lexical-semantic or particularly conceptual representations need to be accessible even after processing of a sentence is completed.

Thus, the model predicts that a recall contribution of lexical-semantic or conceptual information should still be observable after intervening sentences but that this should not be the case for phonological information. The model provides a framework within which short-term retention of different types of information can be investigated. By including assumptions about activation duration, it also predicts when a certain type of information should no longer be retained. This specification is not included in Figure 8 above and can thus not be concluded from it. Here, one node and one link represent more than one kind of representation or process. For instance, the morphosyntactic node represents different features such as grammatical gender or subcategorization features with possibly different retention characteristics.

For most of the representational levels postulated in the model, an involvement in sentence recall has already been found. I will briefly review some of the evidence in the following paragraph. Nearly all of the studies reported here have applied Potter and Lombardi's (1990) intrusion paradigm or modifications of that paradigm, which has also been the case in the present study.

As described above, evidence for the contribution of conceptual and lexical-semantic information comes from the first study by Potter and Lombardi (1990). The question of whether a contribution of syntactic information can be observed has been addressed by Potter and Lombardi (1998) and by Rummer, Engelkamp, and Konieczny (2003). Potter and Lombardi found that the syntactic structure of the second clause of a two-clause sentence influenced recall of the first clause. That is, processing of a not yet recalled clause influenced the syntactic structures used in sentence recall. This is in line with observations of syntactic priming in 'normal' sentence production (see above). Rummer et al. (2003) observed a more direct influence of syntactic structure in terms of recall differences for two-clause sentences with either

subordinate or coordinate structures. Verbatim recall was more accurate for sentences consisting of a main clause and a subordinate clause than for sentences consisting of two main clauses. This finding was attributed to better integration of the subordinate structures into a common structure. The recall difference engaged only the different syntactic structures, since, except for the connective, the wording of the sentences was the same in subordinate and coordinate sentences.

Evidence for the contribution of phonological information comes from Park and Martin (2002) and from Rummer, Martin, Schweppe, and Bormann (subm.), for instance. In a modified version of the intrusion paradigm, they contrasted lure words that were only semantically related to the target word (as was the case in Potter and Lombardi's, 1990, experiments) with lure words that were additionally phonologically related to the target word. They found that participants substituted the target words more often with the lure words that bore both a semantic and a phonological relationship to the target than with those that were only similar in meaning but not in form.

Additional evidence that this increase in intrusion errors involves the storage of phonological information comes from a patient study. IS is assumed to suffer from a selective phonological short-term memory deficit. This implies that she shows a reduced short-term memory span and better performance in a category probe task, which taps semantic short-term memory capacity, than in a rhyme probe task, which taps phonological short-term memory capacity. IS was presented with similar materials as the unimpaired participants. In contrast to the unimpaired participants, she did not show the increased intrusion effect with lure words that were also phonologically related. In contrast, the intrusion effect was numerically even higher for the solely semantically related lure words than for the semantically and phonologically related lure words (Rummer et al., subm.).

In a further sentence recall study, Alloway (subm.) found that purely phonologically related lure words can also induce intrusion errors. Using simply phonological relations between lure and target word requires under-specified sentences that carry hardly any meaning. However, although the influence of phonological lures is rather small, the study by Alloway corroborates previous evidence for the involvement of phonological information in sentence recall.

One central aspect of the model is its modality-specificity. Rummer, Schweppe, and Martin (subm.) observed lower intrusion rates when input and output modality were congruent (i.e., in oral recall of auditorily presented sentences and in written recall of visually presented sentences) as compared to conditions in which they were incongruent (i.e., in oral recall of visually presented sentences and in written recall of auditorily presented sentences). They interpreted this attenuated occurrence of intrusions as reflecting the privileged access to modality-specific linguistic (i.e., phonemic or graphemic) representations.

A different modality influence arises from the contribution of pre-linguistic acoustic-sensory information. For the last one or two items of an auditorily presented list (e.g., Conrad & Hull, 1968; Cowan, 1984; Crowder, 1967; Penney, 1989) or sentence (Balota, Cowan, & Engle, 1990) a sound representation can be available. When used, the sound representation improves recall performance as compared to visual presentation. In contrast to the advantage due to a congruent input-output relation, the use of this sound representation improves both oral and written recall of auditorily presented materials (Penney, 1979). In the intrusion paradigm, the contribution of acoustic-sensory information has been demonstrated by Rummer and Schweppe (2005). All of these surface representations can help to prevent intrusions and thus reduce the impact of lexical priming. Intrusion effects are stronger in conditions that do not provide these representations or in which they cannot be accessed.

With respect to morphosyntactic information, the only study I know that has investigated a recall contribution so far has found no influence (Lombardi & Potter, 1992, see also above). However, according to Rummer and Schweppe's model, whether or not a recall contribution can be expected depends on the *type* of morphosyntactic information. Following from the assumption that requirements of language processing determine what representations are retained, a short-term memory influence of morphosyntactic information should be restricted to such morphosyntactic representations that are relevant at least until a sentence has been fully interpreted. This is not clearly the case for the morphosyntactic feature that Lombardi and Potter (1992) have investigated (see also Subchapter 1.4.2 above). Whether an object occurs in the form of a noun phrase or of a prepositional phrase usually makes no difference in meaning and a representation of the surface form is not necessary for understanding a subsequent sentence. Therefore, it is in line with the predictions that

Lombardi and Potter (1992) did not find a recall influence of the subcategorization features of verbs. For more relevant morphosyntactic features the picture should be different.

In German (as well as in various other languages, though not in English), information about number, case, or grammatical gender of nouns is relevant for processing operations such as thematic role assignment (“Who did what to whom?”, etc.) or for recognizing and establishing syntactic agreement between related elements within or between sentences and phrases. In the following German examples, a change in thematic roles arises from a case change of the determiners only (7a and 7b) and the agent of the second clause changes solely because of the different grammatical gender of the pronoun (8a and 8b). Both variations change the sentences’ meanings substantially, in example (7) even the verb meaning changes. Literal translations and – where necessary – glosses of the examples are given in brackets.

(7a) Ich schenke den_[accusative] Schlumpf_[accusative] einem_[dative] Kind_[dative].

Literal translation: I donate the_[accusative] smurf_[accusative] a_[dative] child_[dative].

Gloss: I give the smurf to a child.

(7b) Ich schenke dem_[dative] Schlumpf_[dative] ein_[accusative] Kind_[accusative].

Literal translation: I donate the_[dative] smurf_[dative] a_[accusative] child_[accusative].

Gloss: The smurf and I are having a child.

(8a) Der_[masc] Bruder schlägt die_[fem] Schwester, weil er_[masc] böse ist.

Literal translation: The_[masc] brother beats the_[fem] sister because he_[masc] mean is.

Gloss: The brother beats the sister because he is mean.

(8b) Der_[masc] Bruder schlägt die_[fem] Schwester, weil sie_[fem] böse ist.

Literal translation: The_[masc] brother beats the_[fem] sister because she_[fem] mean is.

Gloss: The brother beats the sister because she is mean.

Contrary to the morphosyntactic features investigated by Lombardi and Potter (1992), these types of morphosyntactic information are (at least potentially) relevant

for *comprehending* the sentence and for integrating incoming information and thus need to be maintained for some time. For instance, grammatical gender information helps to identify the correct antecedent of a pronoun, which requires retention of the gender of all possible antecedents. According to Garnham, Oakhill, Ehrlich, and Carreiras (1995, p. 49), “the use of gender cue is an integral part of understanding sentences, even when the gender is purely grammatical and would not naturally be considered part of a representation of content” (see also Chapter 2.1 below).

As an additional hint that grammatical gender might play a similar role in reproduction of sentences as it plays in production of sentences, data from a French neuropsychological patient, CB, can be cited. CB shows a selective deficit in written production of closed-class words in sentence contexts. He has no problems in oral production or in producing closed-class words in isolation but shows comparable deficits in writing to dictation (i.e., in written sentence reproduction). Most affected are determiners and most errors on determiners concern grammatical gender while determiner form (definite or indefinite) and number are largely preserved (Alario & Cohen, 2004). The finding that this error pattern occurs in both production and reproduction of sentences gives reason to assume that grammatical gender representations are indeed shared between sentence processing and sentence retention.

So in spite of Lombardi and Potter’s (1992) results, Rummer and Schweppe’s model of short-term memory for syntactically structured materials would predict that ‘more influential’ morphosyntactic information such as grammatical gender can contribute to sentence recall beyond syntactic priming and that thus a more direct influence should be observable.

To my knowledge, apart from Lombardi and Potter’s (1992) study, morphosyntactic representations have not yet been taken into account in short-term memory research. This is probably due to a focus on retention of lists, which in the cases of digits, letters, or non-words obviously do not offer morphosyntactic information. And in the case of word lists, the relevance of such information can be assumed to be rather negligible. Nevertheless, a contribution of morphosyntactic information to short-term retention might well be observed when more complex materials than unrelated lists are investigated.

To address this question, I have examined short-term recall of sentences as did Lombardi and Potter (1992), and, as an extension of this, I have also increased

the complexity of materials to short text passages. Sentences (and texts) provide *propositional-conceptual*, *syntactic* and *morphosyntactic* information in addition to *lexical-semantic* and *phonological* representations that are also provided by single words. As a characteristic and important morphosyntactic feature in German, I have selected grammatical gender to be investigated with respect to short-term memory for morphosyntactic information. In the following chapter, I will therefore briefly describe the (processing) characteristics of grammatical gender in German and will review relevant studies on how grammatical gender information influences language perception and language production.

2. Grammatical gender

The present investigation of possible recall influences of morphosyntactic information focuses on *grammatical gender*, which is a characteristic property of nouns in many languages. German, for example, has a three-gender-system (masculine, feminine, and neuter), in which gender is specified on determiners or adjectives but not on the nouns themselves and in which there is often no clear correspondence between the grammatical and the conceptual gender of a noun. This can be illustrated by having a look at cutlery: In German, a fork is syntactically feminine (“die_[fem] Gabel_[fem]”), a spoon is male (“der_[masc] Löffel_[masc]”), and a knife is neuter (“das_[neut] Messer_[neut]”). Obviously, there is no natural gender difference for types of cutlery (and even an anthropomorphic interpretation would suggest different gender attributions). While other types of morphosyntactic features such as case or number often have clear reflections in conceptual features (e.g., “two husbands” mean something else than “one husband”), grammatical gender (in German) provides a good means for investigating morphosyntactic influences while minimizing the influence of lexical semantics.

The following example (9) (adapted from Bosch, 1988) illustrates a situation in which one and the same object can be referred to with nouns of different grammatical genders. Moreover, referring pronouns each agree in grammatical gender with the *last-mentioned* noun phrase. This suggests that it is necessary to maintain a representation in which information about grammatical gender and meaning of a noun is linked and to update this memory record as soon as a new noun denoting the same referent is introduced.

(9) Als ich *das Auto*_[neut] kaufte, war *es*_[neut] erst ein Jahr alt. Dann habe ich *den Wagen*_[masc] fünf oder sechs Jahre gefahren und eigentlich nie Ärger mit *ihm*_[masc] gehabt. Voriges Jahr schließlich habe ich *die alte Kiste*_[fem] dann verkauft, weil *sie*_[fem] überall anfang zu rosten.

Literal translation: When I *the auto*_[neut] bought, was *it*_[neut] only a year old.

Then have I *the car*_[masc] five or six years driven and really never trouble with

*it*_[masc] had. Last year eventually have I *the old rattletrap*_[fem] then sold because *it*_[fem] everywhere started to rust.

Gloss: When I bought *the auto*, *it* was only a year old. Then I drove *the car* for some five or six years and really never had any trouble with *it*. Last year eventually I sold *the old rattletrap* because *it* started to rust everywhere.

What is more, grammatical gender has already been thoroughly investigated in the context of language processing and serves as a means of testing models of lexical access (in both perception and production). It has been found to have an influence on various tasks, among them lexical decision, picture and word naming. Beyond lexical access, grammatical gender also plays a prominent role in antecedent identification in discourse processing. Gender congruency to a pronoun along with semantic plausibility is an important factor in determining the noun (phrase) the pronoun refers to.

In the following subchapters, I will review some of the studies that demonstrate the relevance of grammatical gender: first, in the context of language comprehension, including studies on lexical access and on anaphor processing (Subchapter 2.1) and, second, in the context of language production, focussing on gender effects on selection of nouns (Subchapter 2.2). Taken together, these studies give reason to assume that gender is in fact a promising morphosyntactic feature to be investigated in a short-term memory context.

2.1 Grammatical gender in language comprehension

With respect to language comprehension, grammatical gender has often been investigated in the context of access to lexical entries and in the context of pronoun interpretation. In both lexical decision and word naming tasks, it has been demonstrated that prior presentation of gender-specified words (determiners, adjectives, etc.) can influence lexical access in two directions. If the modifier has a different grammatical gender than the target noun (i.e., if they are gender incongruent), participants react more slowly than after a neutral prime. These inhibitory effects are quite robust and have been observed in various languages (e.g., in German: Schmidt, 1986; Jacobsen, 1999; in French: Grosjean, Dommergues, Cornu, Guillelmon, & Besson, 1994; in Italian: Bates, Devescovi, Hernandez, & Pizzamiglio, 1996).

Facilitative effects seem to depend more strongly on the experimental conditions. Sometimes faster reaction times if the modifier and the target noun have the same grammatical gender (i.e., if they are gender congruent) have been found (in particular with auditory presentation, in French: Grosjean et al., 1994; in Italian: Bates et al., 1996; in Russian: Akhutina, Kurgansky, Polinsky, & Bates, 1999). Sometimes faster reaction times have not been found (in particular with visual presentation, e.g. in German: Jacobsen, 1999; for an overview of gender congruency effects in language comprehension see Friederici & Jacobsen, 1999). Although it is not clear whether gender information does indeed speed up lexical access, it can be concluded that readers (or listeners) process grammatical gender of modifiers such as determiners or gender-specified adjectives immediately and that the processing of grammatical gender influences the search process in the mental lexicon.

In addition to these gender priming effects, Schriefers, Friederici, and Rose (1998) demonstrated that a gender mismatch between prime and target can even eliminate semantic priming effects. They presented sentences in which the semantic relation between verb and target noun (on which a lexical decision had to be performed) was varied (e.g., for a strong semantic relation "Er schreibt den Brief_[masc] /He writes the letter]"). Additionally, they manipulated whether the determiner agreed or disagreed with the noun. Semantic priming effects of the verb on the noun were obtained when the noun and its preceding determiner agreed in grammatical gender but could not be observed when they disagreed.

Wicha, Orozco-Figueroa, Reyes, Hernandez, Gavaldón de Barreto, and Bates (2005) embedded a picture naming task in sentence comprehension and manipulated the gender congruency of a determiner preceding the to-be-named picture and the semantic adequacy of the picture in the sentence. Participants named the pictures more slowly after a gender incongruent than after a gender congruent prime and named semantically incongruous pictures more slowly than semantically congruous ones. Responses were slowest when both semantic and morphosyntactic cues were incongruent with the picture. These results, in particular the latter two studies, indicate the great impact of morphosyntactic information on lexical-semantic processing and clearly suggest interactions between both kinds of processes.

Beyond word and sentence boundaries, a processing influence of grammatical gender is also observable. Grammatical gender can be used to check which discourse

entities agree syntactically with an anaphoric expression (such as a pronoun). When referring to a noun mentioned just before, one normally does not repeat it, but uses referring expressions such as anaphoric pronouns (see example (10) below).

(10) Ich beobachtete ein Huhn_[neut], das einen Wurm_[masc] jagte. Plötzlich kam *es*_[neut] auf mich zu.

Literal translation: I watched a chicken_[neut] that a worm_[masc] chased. Suddenly came *it*_[neut] towards me.

Gloss: I was watching a chicken that was chasing a worm. Suddenly it approached me.

In the above example, two nouns (i.e., “Huhn” and “Wurm”) are possible antecedents for the anaphoric pronoun “es”. Various information sources can be used to resolve an anaphor. Often the antecedent was part of the last clause, has been in the focus or has had the same grammatical role as the anaphor (e.g., Sanford & Garrod, 1981). In addition to these criteria, morphosyntactic agreement is an important factor. In a language with grammatical gender, detecting which noun is being referred to is – among other factors – guided by gender agreement between the anaphoric expression and possible referents. The fact that the noun phrase “einen Wurm” is masculine and “ein Huhn” is neuter links the neuter pronoun “es” to “ein Huhn”. Meyer and Bock (1999, p. 289) suggest that “speakers try to access the grammatical features of lexical items that were previously used to denote the same referent in the active discourse context.”

The role of gender information in this process has been demonstrated in several studies. The most decisive evidence comes from studies with non-English participants since due to the English gender system, *grammatical* and *conceptual* gender are confounded in English experiments (they compare male, female, and inanimate antecedents). Thus, these studies cannot be interpreted as reflecting the influence of morphosyntactic representations alone. However in languages with a grammatical gender system, grammatical gender supports anaphor interpretation independent of additional conceptual gender marking.

Garnham et al. (1995) investigated how grammatical and conceptual gender influenced antecedent identification for pronouns in Spanish and French. Pronouns

whose antecedents could be identified unambiguously on the basis of gender information excelled pronouns without ambiguous gender information in both study languages and in various measures: they were read faster and questions that required correct anaphor resolution were answered more quickly and more accurately when grammatical gender information allowed for antecedent identification. These results occurred independently of whether the antecedent was animate (and gender was thus specified conceptually) or inanimate (and gender was thus specified only morpho-syntactically).

In their final (French) experiment, Garnham et al. (1995) increased the distance between antecedent and pronoun. With one more clause and a sentence boundary intervening, gender uniqueness still provided advantages with respect to all dependent variables but the benefit was larger when the pronouns referred to persons than to objects. Garnham et al.'s (1995) results provide strong evidence for the influence of grammatical gender information in anaphor interpretation, which is still observable but reduced with greater distance between antecedent and anaphoric pronoun.

These findings from language comprehension indicate that one makes use of gender cues (such as gender-specified determiners) to identify correct entries in the mental lexicon. And they indicate that processing is significantly inhibited by a mismatch between the gender information provided by a cue and the semantic information provided by the context. In addition, the results concerning anaphor interpretation suggest that the cuing influence of gender information does not only work forwards but also backwards. That is, gender cues do not only gate subsequent word processing but provide hints as to which previously encountered word a pronoun might refer to. In the following paragraphs, I will turn to the influences gender information exerts on language *production*.

2.2 Grammatical gender in language production

In language production, a vast amount of the data on grammatical gender effects stems from picture naming studies. Schriefers (1993) obtained a gender congruency effect in the production of gender-specified noun phrases in Dutch using a modified version of the picture-word interference paradigm. In this paradigm, participants have to name a picture while ignoring a distractor word that is presented either at the

same time as the picture or shortly before or after picture presentation. The effect this distractor word has on picture naming can enlighten the (sub-) processes involved in word production and its time course. For instance, semantically related words presented briefly before the picture delay speech onset latencies. This indicates early competition at the semantic level. In contrast, phonologically related distractors speed up the production process – if they are presented briefly after the picture. This indicates late activation of phonemes because phonemes that are shared between target and distractor receive activation from both sources (Schriefers, Meyer, & Levelt, 1990).

Schriefers (1993) had his participants name pictures using noun phrases consisting of a determiner, a color adjective, and a noun. The distractor noun was either of the same gender as the target noun or of different gender. When the picture and the distractor word were presented simultaneously, participants started naming the picture later with a gender incongruent noun than with a gender congruent noun (*gender congruency effect*). Gender features are assumed to be linked to lexical representations, either to a lemma (a kind of grammatical word that assembles morpho-syntactic representations, see e.g., Levelt, 1989; Levelt et al., 1999) or directly to the phonological word form (see e.g., Miozzo & Caramazza, 1999). Schriefers (1993) explained the gender congruency effect as resulting from interference between gender nodes which again results in delayed speech onset if target and distractor are of different genders.

This effect has been replicated for Dutch by van Berkum (1997), La Heij, Mak, Sander, and Willeboordse (1998), and Schiller and Caramazza (2003, Exp. 2) and for German by Schriefers and Teruel (2000) and Schiller and Caramazza (2003, Exp. 1). For Italian, Miozzo and Caramazza (1999) did not find a gender congruency effect (but see Cubelli, Lotto, Paolieri, Girelli, & Job, 2005, for a gender influence on Italian picture naming), nor did Costa, Sebastián-Gallés, Miozzo, and Caramazza (1999) for Catalan and Spanish.

Miozzo and Caramazza (1999) explained these contrasting results in terms of a structural distinction between the respective languages concerning the time course of determiner selection. In both Dutch and German, selection of a gender-specified determiner depends solely on the gender of the adjacent noun (no matter what the phonological form of the corresponding noun, the definite determiners in German are

always *der* for masculine, *die* for feminine, and *das* for neuter nouns; in Dutch *het* belongs to neuter and *de* to common gender). Therefore, Miozzo and Caramazza (1999) termed these languages *early-selection languages*. In contrast, in Italian, for instance, determiner selection depends not only on the gender of the noun, but also on the phonology of the following word (i.e., the onset in this case). The singular definite determiner for a masculine noun in Italian can either be *lo* (preceding words that start with either a vowel, with an s + consonant, with gn, or with an affricate; e.g., *lo zucchero* – the sugar) or *il* in all other singular cases (e.g., *il caffè* – the coffee). Hence, determiner selection cannot occur until the subsequent phonological context has been specified. This kind of language is called a *late-selection language*. Different determiner selection procedures can thus account for different consequences of gender incongruent distractors on noun phrase production (for an overview of cross-linguistic differences on gender congruency effects see Caramazza, Miozzo, Costa, Schiller, & Alario, 2001). With respect to the present study, this distinction is relevant only for the classification of the study language. All of the experiments that will be reported here have been conducted in German, which is an early-selection language. In early-selection languages, gender congruency effects have consistently been observed in the production of noun phrases. This provides a good basis for the potential observation of a gender congruency effect in the reproduction of noun phrases.

However, in addition to the type of language, there are other conditions that constrain the occurrence of gender congruency effects in picture-word interference tasks. The gender congruency effect seems to be restricted to cases in which not only the gender features, but also the gender-specified determiners of target and distractor differ. La Heij et al. (1998) failed to replicate the gender congruency effect for the production of Dutch bare nouns and Schiller and Caramazza (2003) did not find larger interference from gender incongruent distractors if participants had to produce German plural noun phrases.²

² Note that in German (as well as in Dutch) different gender types are associated with different definite determiners only in singular form (*der* for masculine, *die* for feminine, and *das* for neuter), whereas all definite plural determiners are of the same form (*die*). Thus, in the case of plural noun phrases consisting of a determiner and a noun, gender incongruency does not imply determiner incongruency.

The gender congruency effects reported so far consist in differences in onset latencies. More closely related to the present study are *error* patterns reflecting an influence of grammatical gender. Semantic substitutions in German slips of the tongue tend to preserve the grammatical gender of the target noun more often than is predicted by chance (Marx, 1999), as do aphasic misnamings in German (Kulke & Blanken, 2001). An influence of grammatical gender on errors also holds for experimentally induced semantic substitution errors.

Vigliocco, Vinson, Indefrey, Levelt, and Hellwig (2004) had their German-speaking participants name rapidly presented pictures which were presented in blocks consisting of one semantic field (e.g., animals) in order to elicit semantic errors. When semantic and phonological similarity were partialled out, grammatical gender significantly predicted errors when the task required production of gender-specified phrases (gender-specified determiner plus noun) but did not when the pictures had to be named by bare nouns or nouns in combination with determiners underspecified for gender. The authors concluded that grammatical gender influences semantic errors only in those cases in which syntactic frames requiring gender information are present and in which gender information is also specified morphophonologically.

Both studies that have investigated response latencies and studies that have focused on a gender influence on spontaneous or induced speech errors have observed a gender influence on lexical selection. It is important to note that this influence was solely observed when nouns had to be produced in a (minimal) syntactic context. This difference between bare noun production and production of syntactically structured utterances can also give hints for the investigation of the role of grammatical gender information in retention of verbal materials. If grammatical gender does not contribute significantly to bare noun production, why should it make a significant contribution to bare noun reproduction? As reported above, serial recall of unrelated word lists provides the opportunity to systematically vary several characteristics of the materials to be recalled. Studies on verbal short-term memory that have applied this paradigm have demonstrated recall influences of phonological factors (phonological similarity and word length), lexical factors (lexicality and word frequency), and semantic factors (concreteness and imageability). Yet the above-

reported language production studies suggest that a recall influence of grammatical gender should only be observable as far as syntactically structured materials are used.

In sum, three main lessons can be learned from the studies that have been sketched in this chapter: (1) grammatical gender of nouns influences both language comprehension and language production, (2) grammatical gender information needs to be *maintained* for successful discourse comprehension, and (3) observing an influence of grammatical gender requires the use of syntactically structured materials.

So far, Chapter 1 has outlined the working memory models on which the present study is based (this is particularly the case for the one by Rummer & Schweppe) and Chapter 2 has reported findings concerning the role of grammatical gender in language comprehension and language production. Based on this, Chapter 3 will recapitulate what can be inferred for the present study. It is meant to outline what the study aims at and why. In addition, it will provide a sketch of the main paradigm that I have applied to address the question of shared representations in language processing and verbal short-term memory by the example of grammatical gender.

3. Objectives of the present study

Chapter 3 is meant to provide an interface between the theoretical foundations described above and the experimental approach to this study's basic research question, which will be described below. The present study addresses the question whether retention of verbal materials relies on the same representations as language processing, which is the basic assumption that underlies language-based models of short-term memory. Specifically, I want to investigate whether grammatical gender information, which is needed for sentence and text processing, exerts an influence on recall of syntactically structured materials. Hence, I have sketched models of (verbal) working memory and have given an overview of the role of grammatical gender in language processing in the previous chapters. As indicated above, the model proposed by Rummer and Schweppe (in prep.) serves as my main theoretical foundation. Therefore, the hypotheses for the following experiments will be deduced on the basis of this model. However, I will also report assumptions based on Baddeley's (2000) and Cowan's (1999) model at the beginning of this chapter. As will be outlined below, these models are less precise in their predictions for the questions at hand. Thus, their implications concerning the role of grammatical gender in short-term recall of syntactically structured materials will be described only briefly. Following the assumptions based on language-based models of short-term memory in general and on Rummer and Schweppe's model specifically, I will describe the paradigm that was the basis of the experiments reported here in the final part of this chapter.

Predictions on the basis of Baddeley's (2000) working memory model. In Chapter 1, different theories of verbal short-term memory have been outlined. These theories differ in the way they assume language processing and short-term memory for linguistic materials to be intertwined. Baddeley (2000) has clearly postulated separate short-term memory components that can be filled with linguistic long-term knowledge (the episodic buffer) or with surface representations of the to-be-recalled verbal materials (the phonological loop). In Baddeley's (2000) model, representations from the language processing system only come into play when they get into the episodic buffer. The inclusion of the episodic buffer in the original model demonstrates the need for linguistic representations to take part in retention of verbal materials. How-

ever, the model does not specify what types of representations are involved in sentence recall and how they might interact. Therefore, the model does not make any predictions with respect to the involvement of morphosyntactic representations in recall and cannot be consulted for deriving hypotheses.

Predictions on the basis of Cowan's (1999, 2001) memory model. Cowan's model links short-term memory more closely to long-term memory. Rather than assuming separate stores, Cowan differentiates separate degrees of activation and awareness. With respect to sentence repetition, he states that "if we attempt to repeat a sentence, we do not repeat the acoustic waveform; we determine the known units that correspond to what was said and then attempt to produce those units, subject to the capacity limit (Cowan, 2001, p. 92)." Thus, in the context of the present study, his model provides a more general framework for the language-based models that have been presented. This is particularly the case for the one proposed by Dell et al. (1997), which conceives of verbal short-term memory as prolonged activation of representations in the language processing system, and for the conceptual regeneration hypothesis put forward by Potter and Lombardi (1990). Implications of Cowan's model can thereby be incorporated in the discussion of these theories.

Predictions on the basis of language-based models of short-term memory. Language-based models of short-term memory have motivated the present study. The contribution of both phonological and lexical-semantic information to short-term memory tasks is now widely accepted. This study aims at extending evidence for language-based models by investigating a type of information that can clearly be regarded as relevant in language processing and that would not be assumed to play a role in short-term memory tasks unless from a language-based point of view: namely, grammatical gender information.

As indicated above, grammatical gender plays an important role in processing of syntactically structured materials and should thus also play a role in their retention. However, the dominant paradigm in verbal short-term memory research has been and still is immediate serial *list* recall (e.g., Eysenck & Keane, 2005). The working memory model proposed by Baddeley relies to a large degree on serial list recall data. On the basis of this paradigm, it has been demonstrated that the modality

in which materials are presented influences how accurately they can be recalled and that word length, phonological similarity, word frequency, and imageability of the items on a list influence how many items can be recalled correctly. Also focusing on list retention, research on parallels between language processing and verbal short-term memory has mostly concentrated on the recall contribution of phonological and lexical-semantic representations (Dell et al., 1997; Martin et al., 1999). As evidence for the assumed close link between word processing and word retention, patient data and data that demonstrate the involvement of linguistic representations in short-term memory tasks are cited.

An influence of grammatical gender information in language processing has been observed with (at least minimally) syntactically structured materials but not consistently with bare nouns. If an influence in a short-term memory task can be observed, this should also be the case in retention of syntactically structured materials. As it is postulated that retention and processing of verbal materials rely on the same representations, the conditions under which grammatical gender exerts an influence should be similar in short-term memory and language processing tasks. A principal characteristic of grammatical gender in language processing is that it “marks as equal” two or more related elements. These elements can be adjacent (as an adjective and a following noun, which have to agree in grammatical gender, e.g., “kleines_[neut] Mädchen_[neut]” (little girl) vs. “kleiner_[masc] Junge_[masc]” (little boy)) or they can be separated by other words (as a noun and an anaphoric pronoun, which also have to agree in grammatical gender, e.g. “Das_[neut] Mädchen_[neut] entdeckt einen_[masc] Käfer_[masc]. Es_[neut] freut sich.” (*literal translation*: The_[neut] girl_[neut] discovers a_[masc] beetle_[masc]. It_[neut] is happy; *gloss*: The girl discovers a beetle. She is happy.)). This suggests that the potential contribution of grammatical gender to retention of verbal materials should also be most relevant in retention of syntactically structured materials and not in retention of lists that consist of a few single nouns that are not related to one another.

The use of syntactically structured verbal materials instead of unrelated lists has additional advantages. As indicated in the introductory chapter, the role of verbal working memory is often considered with respect to tasks that tap complex cognitive skills. However, these tasks – such as language comprehension or reasoning – usu-

ally require processing and storage of more complex materials than single words or lists of unrelated single words. Solving problems or drawing conclusions involves processing of at least sentences if not texts. On the basis of the studies investigating retention of sentences that I have reported above, it is questionable whether working memory or short-term memory tasks that are based on retention of unrelated materials are an adequate approximation for the kind of working memory that one can assume to be involved in higher-order cognition tasks. For these two main reasons (first, the fact that grammatical gender influences are tied to syntactic contexts and, second, the increased applicability of the results to more complex cognitive tasks), I will use syntactically structured verbal materials in my experiments.

Some theories of verbal short-term memory do already focus on retention of syntactically structured materials and claim that there are important differences between recall of unrelated verbal materials and recall of syntactically and conceptually structured verbal materials. In line with this assumption is the general observation that, depending on the task, different types of representations contribute to recall to different degrees. Roughly speaking, list recall is particularly dominated by phonological information with supporting influence of semantic information. In sentence recall, the influence of phonological information is smaller than the influence of semantic representations (e.g., Jefferies et al., 2004). In addition, propositional and syntactic information come into play in sentence recall, but not in list recall (Potter & Lombardi, 1990, 1998).

The model on which the present study is based focuses on retention of sentences and assumes a causal relationship between the requirements of language processing and retention characteristics of linguistic representations. According to this account, whether and how long a certain type of information is retained depends on whether and how long this information might be needed for understanding or producing the following utterances. In other words, verbal short-term memory is a function of language processing. Certain representations that are needed to understand subsequent utterances should also be available in short-term memory for a longer time. This applies in particular to meaning-related information. Representations that are dispensable for subsequent language processing should be available in short-term memory only for a short period of time. This should be the case for phonological information, for instance. As indicated above, this basic idea requires the study of

larger discourse entities, since linking short-term memory to the requirements of language processing calls for contexts in which more complex and thus more demanding language processing takes place. Word processing is only a very limited field and naturally does not provide and require as much information as sentence or text processing.

Based on its role in initial language comprehension, in comprehension of subsequently encountered information, and its influence on lexical selection in language production, it is predicted that grammatical gender information should also be retained in short-term memory for syntactically structured materials. Testing this prediction, that is, investigating whether such an influence can indeed be observed is the primary objective of the present study.

The observation of a gender congruency effect in a short-term memory task would make a strong case both for language-based perspectives on verbal short-term memory and for theories that focus on retention of more complex materials. In addition, the importance of grammatical gender for anaphor resolution suggests that it is needed even if an additional sentence follows a critical noun phrase. Thus, the potential influence of gender information should not be restricted to recall of a sentence just processed. So in addition to investigating an influence of grammatical gender information on sentence recall in general, a further purpose of this study is to investigate whether a possible recall influence of grammatical gender is restricted to recall of the most recently encountered sentence or whether it extends to less recently processed sentences.

If gender information does indeed contribute to short-term recall of sentences, in what respect could such an influence be observed? Psycholinguistic studies have demonstrated that grammatical gender principally influences language processing by influencing lexical selection (see also Chapter 2 above). Therefore, an influence on sentence recall should also be observable with respect to lexical selection.

As described above, the intrusion paradigm introduced by Potter and Lombardi (1990) interferes with lexical (re-)selection and applies sentence recall. By combining recall of syntactically structured materials and a manipulation that directly addresses lexical selection the intrusion paradigm should thus serve as an appropriate means for identifying the potential contribution of grammatical gender information to a short-term memory task.

A modification of the intrusion paradigm. To investigate whether grammatical gender contributes to verbatim sentence recall, I have modified Potter and Lombardi's (1990) intrusion paradigm so that it includes a gender manipulation. This manipulation was achieved by varying the gender relation between the target word in the sentence and the lure word on the list. In order to allow for a direct comparison of the errors induced by a gender incongruent and by a gender congruent lure word, three lure conditions have been introduced: first, a control condition, in which the word list contained only words that were not related to the target noun, second, a gender incongruent lure condition, in which a lure word was presented that was semantically similar to the target but of a different grammatical gender, and, third, a gender congruent lure condition, in which a lure word was presented that was semantically similar to the target and had the same grammatical gender (see example (11) below, the target word is italicized and English glosses are given in brackets).³

- (11) Sentence: Dem genervten Pförtner war die *Freude*_[fem] an seinem Beruf durch die arroganten Gäste endgültig verdorben.

(The distressed porter had finally lost *delight* in his job because of the arrogant guests.)

Gender Congruent Lure Word: *Lust*_[fem] (pleasure)

Gender Incongruent Lure Word: *Spaß*_[masc] (joy)

As outlined above, the occurrence of lure intrusions in recall when the lure word has been presented on a separate word list reflects the involvement of propositional-conceptual and lexical-semantic information. If additional surface information for the words in the sentence (such as phonological or acoustic-sensory information) is available during recall, the intrusion effect (i.e., the increase in intrusion errors due to the presentation of the lure word) should be attenuated as compared to conditions in

³ Note that the translations for the target word and the two types of lure words can only be approximations. As there is obviously no gender difference for these words in English, the suggested translations only serve to indicate that more than one word fits the sentence content better than the target word (although the English synonym triplets (target word, gender congruent lure word, and gender incongruent lure word) are not equivalent to the German synonym triplets.)

which no or less surface information is available. Similar effects are expected if gender information does in fact contribute to short-term recall of sentences. Then, the gender incongruity between target word and lure word should aid differentiation between target and lure representations and should hence reduce the intrusion rate of the gender incongruent as compared to the gender congruent lure words. If it can be ensured that both types of lure words do not differ with respect to conceptual features, such an attenuated intrusion effect will speak for a short-term memory representation of the grammatical gender of the target noun.

Therefore, the critical dependent variable in this modified version of the intrusion paradigm is whether the target word in the sentence is reproduced and, if it is not, what is produced instead. Four different response types can be distinguished: first, the target word is *reproduced correctly*, second, it is substituted by the gender incongruent lure word (i.e., a *gender incongruent intrusion*), third, it is substituted by the gender congruent lure word (i.e., a *gender congruent intrusion*), or, fourth, neither the target nor one of the lure words is produced (such cases have been termed *target omissions*).

It is predicted that grammatical gender information does indeed contribute to short-term recall of sentences. In the intrusion paradigm, a gender representation of the target word should have an impact on the response types. The crucial assumption is that presentation of a lure word should affect responses differently depending on which lure word (i.e., gender congruent or gender incongruent) was presented. For gender congruent lure words, a strong *intrusion effect* is expected, that is, the relative proportion of correct target reproductions and intrusion errors is predicted to differ as to whether or not the gender congruent lure word was presented. There should be more gender congruent lure intrusions if the gender congruent lure word is presented on the distractor list than if an unrelated word (i.e., a *control word*) is presented instead. The target word should be reproduced correctly more often in the control condition than in the gender congruent lure condition. In contrast, in the gender incongruent case, the gender mismatch should attenuate (or even eliminate) the intrusion effect. Such an attenuated intrusion effect should result in a lower proportion of induced gender incongruent as compared to induced gender congruent intrusions.

Overview of the experiments. The modified intrusion paradigm is more or less the basis of all experiments reported here, with critical variations each. All experiments test whether and under which circumstances and constraints grammatical gender influences recall of syntactically structured verbal materials. The first one tests whether there is an influence at all. In the second and third experiment, it is tested whether or not such an influence involves a contribution of a phonological representation of the target determiner instead of a contribution of morphosyntactic information. In these experiments, the availability of phonological information is systematically reduced. The fourth experiment tests more specifically whether a gender congruency effect in short-term sentence recall is influenced by the form of the target determiner. It compares intrusions in sentences in which the target determiner is unambiguously specified for gender and is thus incompatible with the gender incongruent lure word with intrusions in sentences in which the target determiner is ambiguously specified and is thus also compatible with the gender incongruent lure word. That is, it tests the assumption that the gender congruency effect does not implicate the fact that the target word and the gender incongruent lure word require different determiners, while target and gender congruent lure words have the same determiner. This experiment thus aims to provide evidence against a determiner-based alternative explanation, with regard to both a phonological and a lexical representation of the determiner. In sum, these experiments aim at pinning down the gender congruency effect satisfactorily to the contribution of morphosyntactic information.

The final experiment expands the scope of this study to recall of more complex discourse. Although sentences are more complex and richer materials than single words, they are still restricted. Experiment 5 uses short texts instead of single sentences and tests whether a recall contribution of gender information is influenced by the processing status of the to-be-retained materials. That is, it investigates whether grammatical gender contributes to recall only when the target word is part of the most recent sentence or whether it also contributes when it is part of a less recent sentence.

4. Experiment 1

The first experiment of the present study aimed at testing whether grammatical gender information can contribute to short-term recall of sentences. It was assumed that the influence of gender information should be observed as an influence on lexical selection. In the intrusion paradigm introduced by Potter and Lombardi, the presentation of a lexical competitor on a distractor word list interferes with lexical selection of a target noun in sentence reproduction.

For the purpose of this study, the intrusion paradigm has been modified so that two lexical competitors for the target word exist: one that has the same grammatical gender and one that has a different grammatical gender than the target word (see example (11) above). The occurrence of more induced intrusions (i.e., intrusions if the respective lure word has been presented) than spontaneous intrusions (i.e., intrusions if neither lure word has been presented) would reflect the general influence of the lexical priming manipulation.

If, in addition to lexical information, information about the gender of the target noun is retained in short-term memory, then such an intrusion effect is expected for gender congruent lure words, since they are assumed to be activated both lexically and morphosyntactically (because of the gender match to the target). For gender incongruent lure words the intrusion effect should either be attenuated or be eliminated because of the gender mismatch to the target word. That is, the activation of the lure word being part of the distractor list should be counteracted by the conflicting gender information. In addition, the presentation of the gender congruent lure word should lead to more intrusion errors than the presentation of the gender incongruent lure word because of the contribution of grammatical gender information. Such a result could be labelled a gender congruency effect in short-term sentence recall.

Naturally, the proportion of intrusion errors influences the proportion of correct reproductions of the target word, since these response types are not independent of one another. Therefore, the two types of intrusion effects (for gender congruent and gender incongruent lure words) and the gender congruency effect (more induced gender congruent than induced gender incongruent intrusions) will be considered in relation to the proportion of correctly reproduced target words.

To ensure that the expected difference between gender congruent and gender incongruent intrusion errors goes back to the gender difference, it has to be made sure that the two types of lure words do not differ from each other with respect to the conceptual adequacy in the sentence and with respect to several lexical-semantic and phonological characteristics. The results of the respective pre-tests will be reported in the materials section of this chapter.

The experiment was based on a one-factorial design with three factor levels. It was manipulated within participants whether the distractor list consisted only of unrelated words (*without lure*), whether a *gender incongruent lure word* was included in the list, or whether a *gender congruent lure word* was presented.

4.1 Method

Participants. 24 students of Saarland University participated in the first experiment. The participants were native speakers of German and were either awarded course credit or were paid for participation.

Materials. 30 sentences were constructed in such a way that each sentence included a noun that was not an optimal fit for the content of the sentence (target word). Instead, there were two semantically related words that were a better fit (lure words), one of them of the same grammatical gender as the target word in the sentence and one of different gender. To ensure that the grammatical gender was the only critical difference between the two types of lure words, several factors had to be controlled. The two types of lure words were controlled for semantic similarity to the target word, the conceptual adequacy in the sentence, word frequency, word length, and phonological similarity to the target word. The respective pre-tests will be described in the following paragraphs. Means and standard deviations for target and lure words in all pre-tests are also summed up in Table 1.

Table 1: Means and standard deviations (in brackets) for the target and lure words in the pre-tests for Experiment 1.

	Target Word	Gender Congruent Lure Word	Gender Incongruent Lure Word
Semantic Similarity (1-10)	--	6.5 (<i>SD</i> =1.8)	6.6 (<i>SD</i> =1.6)
Cloze Probability (%)	2.3 (<i>SD</i> =3.4)	27.5 (<i>SD</i> =20.4)	27.5 (<i>SD</i> =22.0)
Word Frequency (per million)	321 (<i>SD</i> =655)	336 (<i>SD</i> =441)	200 (<i>SD</i> =222)
Word Length (Characters)	6.5 (<i>SD</i> =2.5)	5.6 (<i>SD</i> =2.1)	5.9 (<i>SD</i> =1.9)
Word Length (Syllables)	2.1 (<i>SD</i> =0.8)	1.8 (<i>SD</i> =0.6)	1.8 (<i>SD</i> =0.6)
Shared Phonemes (Final Syllable)	--	0.4 (<i>SD</i> =0.7)	0.4 (<i>SD</i> =0.6)

Pre-tests. There is some evidence that gender congruent nouns tend to be more similar in meaning than incongruent nouns (Köpcke & Zubin, 1984). Hence, in a pre-test, 20 native speakers of German (none of whom participated in one of the experiments) rated the semantic similarity between lure and target words on a scale ranging from 1 (no semantic similarity) to 10 (high semantic similarity). On average, the gender incongruent lure words were rated equally similar to the target words (6.6) as the gender congruent lure words (6.5).

In addition, both types of lure words had to fit the propositional-conceptual sentence content equally well. Conceptual fit was measured by a sentence completion task in which 20 more participants completed the sentences with the target determiners and nouns omitted. Participants were asked to carefully read the sentences and fill in the words that first came to mind. On average, the sentences were completed with the incongruent lure words in 27.5 % of the cases, just as often as with

the congruent lure words, whereas the target word was used in only 2.3% of the cases.

Furthermore, the two different types of lure words matched with respect to their word length in characters and in syllables. On average, the incongruent lure words consisted of 5.9 characters and 1.8 syllables and the congruent lure words consisted of 5.6 characters and 1.8 syllables.

Gender congruent and gender incongruent lure words were compared with respect to phonemic overlap with the target word since there might be a correlation between final syllable and gender type (see Hofmann, 2006, for the influence of morphophonological as compared to semantic gender markers in German). The final syllables of target and gender congruent lure words had a mean of 0.3 overlapping phonemes in the same position and of 0.1 more phonemes in different positions, just like the final syllables of target and gender incongruent lure words.

The necessity to match the lure words with respect to the conceptual adequacy and semantic similarity led to word frequency differences between gender congruent and gender incongruent lure words (336 vs. 200 per million, according to the CELEX database). Although the conceptual fit and the similarity to the target word are assumed to be the decisive factors for intrusion errors, it has to be ruled out that – possibly – different effects of the two types of lure words go back to frequency differences. Therefore, additional analyses excluding those six items in which the difference between gender congruent and gender incongruent lure words is highest (resulting in frequency means of 169 per million for the gender congruent lure words) will be reported.

In addition to the sentences, 30 lists of five unrelated words were constructed. In the lure conditions, one of these words was substituted by the gender congruent or the gender incongruent lure word. A complete list of the sentences and lure words can be found in Appendix A.

Procedure. For each participant, a third of the sentences was presented in combination with the lists including a gender congruent lure word, one third was presented with the lists including a gender incongruent lure word, and one third with the lists that included an unrelated control word instead. Across participants, each sentence appeared equally often in each lure condition. The assignment of the lists to one of these conditions was determined at random. In order to avoid primacy and recency

effects, the lure words were never presented in initial or final list positions, but appeared equally often at list positions 2, 3, and 4. For each trial, the position of the lure word (or of the control word) was the same across the three lure conditions.

Each trial started by pressing the space bar of the computer keyboard. Three asterisks then appeared at the center of the screen for 300 ms followed by a blank screen for 350 ms. Then the word list was presented auditorily via loudspeakers. After an interval of 250 ms, in which a row of percentage signs was displayed, the sentence was presented auditorily (at a mean presentation rate of 125 ms per syllable). After sentence presentation, participants had to recall the sentence orally as accurately as possible. In other words, verbatim recall was required. Finally, the probe word was presented auditorily for 500 ms and participants had to decide via key-stroke whether it had been included in the word list or not. The following example (12) illustrates the trial structure of the immediate recall condition with a gender congruent lure – target relation (English glosses are given in brackets).

- (12) List (with Congruent Lure Word): Gruß Fähre Ball *Garten*_[masc] Sprache
 (greeting ferry ball *garden*_[masc] language)
- Sentence Presentation: Sie betraten voller Freude den *Rasen*_[masc]
 um Gänseblümchen und Klee zu pflücken
 (They happily entered the *lawn*_[masc] in order to pick daisies and clover)
- Sentence Recall (with Intrusion): Sie betraten voller Freude den *Garten*_[masc]
 um Gänseblümchen und Klee zu pflücken
 (They happily entered the *garden*_[masc] in order to pick daisies and clover)
- Probe Word: FÄHRE (FERRY)

Before the experiment started, six practice trials were presented. The experiment lasted about 20 minutes. Participants' utterances were recorded and transcribed. The number (and the proportion) of correct reproductions of the target word, of gender incongruent lure intrusions, and of gender congruent lure intrusions served as dependent variables. In addition, the number (and the proportion) of target omissions will be reported. In previous experiments that have applied the intrusion paradigm,

overall recall accuracy has also served as a dependent variable. The manipulation in the present experiment affects only the target word. However, overall recall accuracy will be reported as a descriptive measure in Appendix B in order to allow for comparisons with other experiments.

4.2 Results

Table 2 reports absolute frequencies and percentages of the four different response types (i.e., correct target reproductions, gender incongruent lure intrusions, gender congruent lure intrusions, and target omissions) per lure condition (i.e., list without lure, list with gender incongruent lure, and list with gender congruent lure). In line with the prediction, there was an intrusion effect (i.e., more lure intrusions if the list included a lure word as compared to cases in which the list included a control word instead) for gender congruent lure words but not for gender incongruent ones. Moreover, the proportion of induced intrusions was higher for gender congruent than for gender incongruent lure words, whereas there were hardly any spontaneous intrusions, neither gender congruent (2.1%) nor gender incongruent (0.4%).

Table 2: Absolute frequencies and percentages (in brackets) of gender congruent and gender incongruent lure intrusions, correctly reproduced target words, and target omissions as a function of lure condition in Experiment 1.

	Without Lure	With Gender Congruent Lure	With Gender Incongruent Lure
Gender Congruent Intrusions	5 (2.1%)	24 (10%)	2 (0.8%)
Gender Incongruent Intrusions	1 (0.4%)	3 (1.3%)	3 (1.3%)
Correct Target Reproductions	223 (92.9%)	198 (82.5%)	223 (92.9%)
Target Omissions	11 (4.6%)	15 (6.3%)	12 (5.0%)

Statistical analyses were based on log-linear models (e.g., Howell, 2002), which are appropriate for determining the fit between observed and expected cell counts (as do standard cross-tabulation chi-square tests) and provide the opportunity to test for main effects and interactions in a multi-factorial design. Even more important, log-linear analyses allow for analyzing categorical frequencies, meaning that they do not require factor levels to be independent of one another and do not rely on parametric assumptions concerning the dependent variable (cf. e.g., Scheepers, 2003).

For each analysis I report results when subjects were included as a factor and results when items were included in order to assess the generality of effects across individuals and materials. This implies that expected frequencies for the effects of interest have been adjusted for inter-individual (or inter-item) variation. Likelihood Ratio Chi-Squares (*LRCS*) will be reported with *LRCS1* for the analyses including the subject factor and *LRCS2* for the analyses including items as a factor.

Both partial and marginal associations *LRCS* were computed. Partial associations derive from tests in which expected frequencies are calculated from all effects competing with the respective effect at the same level of the effect hierarchy; marginal associations derive from tests in which expected frequencies are calculated from all effects constituting the respective effect at a more general level of the effect hierarchy. For each effect, I will report only the more conservative term (min. *LRCS*).

To test the hypotheses (particularly the assumption that there are more congruent than incongruent intrusions if the respective lure word was included in the list), three types of analyses are required. Similar to the procedure in standard Chi-square tests, each of these analyses compares intrusion errors and correct reproductions of the target word as a function of lure condition.

(1) The first type of analysis addresses the question whether or not an intrusion effect (i.e., more induced than spontaneous intrusions) can be observed for gender congruent lure words. To do so, two lure conditions are included in the analyses. Gender congruent lure intrusions and correct target reproductions are compared in the condition in which neither lure word was on the list and in the condition in which the gender congruent lure word was on the list. An *intrusion effect for gender congruent lure words* would be reflected in a higher proportion of gender congruent intrusions when the gender congruent lure word was on the list than in the condition in

which a control word was presented instead (i.e., in a two-way interaction between lure condition and response type).

(2) The second type of analysis addresses the same question with respect to gender incongruent lure words. Hence, gender incongruent lure intrusions and correct reproductions of the target word in the control condition (i.e., list without lure word) and in the condition in which the gender incongruent lure word was presented are included in the analyses. As with gender congruent lure words, an *intrusion effect for gender incongruent lure words* could be demonstrated if the proportion of gender incongruent intrusions was higher in the gender incongruent lure condition than in the without lure condition (i.e., if the two-way interaction between lure condition and response type was significant).

(3) Most important is the third type of analysis. It contrasts *induced gender incongruent intrusions* with *induced gender congruent intrusions*. That is, gender incongruent intrusions and correct target reproductions in the condition in which the gender incongruent lure word was presented on the distractor list are compared with gender congruent intrusions and correct reproductions of the target word in the condition in which the gender congruent lure word was presented. In these analyses, a *gender congruency effect* is expected. A gender congruency effect would be indicated by a higher proportion of induced gender congruent intrusions as compared to induced gender incongruent intrusions (i.e., by a two-way interaction between lure condition and response type).

Since my hypotheses concern correct target reproductions and intrusion errors only, target omissions will not be included in the log-linear analyses. As revealed by a Friedman test, there was no significant difference between the three lure conditions (4.6% vs. 5.0% vs. 6.3%, $p > .9$) with respect to target omissions.

(1) *Intrusion effect for gender congruent lure words.* To analyze whether there was a significant intrusion effect for gender congruent lure words, I compared the correct responses and the gender congruent intrusions in those cases in which no lure word had been presented and in those in which the gender congruent lure word occurred on the list. An intrusion effect was expected, that is, more intrusions if the congruent lure word occurred on the list as compared to the without lure condition. This should result in a two-way interaction between lure condition (without lure vs. with congru-

ent lure) and response type (correct target reproduction vs. congruent lure intrusion). As expected, this interaction reached significance (min. $LRCs1=8.30$, $df=1$, $p<.01$; min. $LRCs2=7.54$, $df=1$, $p<.01$).

The interaction was also significant in the analyses excluding six items with substantial frequency differences between gender congruent and gender incongruent lure words (min. $LRCs1=4.11$, $df=1$, $p<.05$; min. $LRCs2=4.11$, $df=1$, $p<.05$). In neither of these analyses the three-way interaction with participants or items was significant (all $ps>.97$). Gender congruent intrusions were more frequent if the lure word had been presented on the distractor list (10%) than if no lure word had been presented (2.1%), whereas the target was correctly reproduced more often in the “without lure” condition (92.9%) than in the “gender congruent lure” condition (82.5%). This reflects an intrusion effect for gender congruent lure words. The results are depicted in Figure 9.

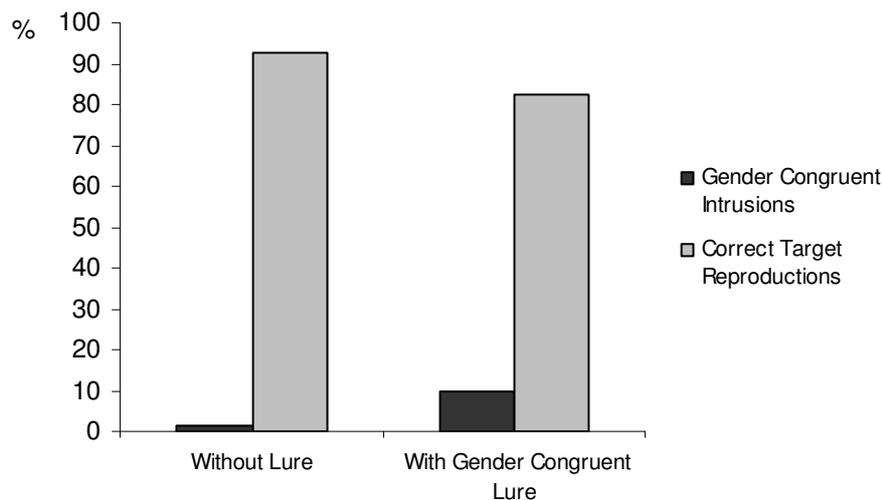


Figure 9: Percentages of gender congruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 1.

(2) *Intrusion effect for gender incongruent lure words.* Equivalent analyses were computed for gender incongruent intrusions (including the “without lure” and the gender incongruent condition and the correct responses and the gender incongruent intrusions). The intrusion effect is assumed to be reduced by the gender mismatch between target and lure word. Thus, a smaller or even no difference between the “with incongruent lure” and the “without lure” condition with incongruent lure intrusions is expected. In line with this expectation, log-linear analyses did not reveal a

significant two-way interaction between lure condition and response type (both $LRCSS < 1$, $df = 1$, both $ps > .7$), which suggests that there was no significant difference between spontaneous (0.4%) and induced intrusions (1.3%) for gender incongruent lure words (see Figure 10 below).

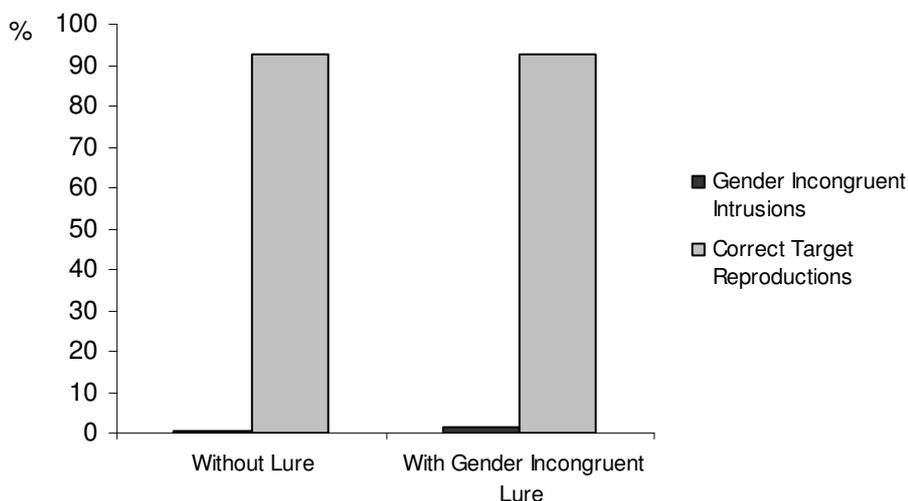


Figure 10: Percentages of gender incongruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 1.

(3) *Induced gender incongruent vs. induced gender congruent intrusions.* In this most critical analysis, correct responses and gender incongruent intrusions in the “incongruent lure” condition and correct responses and gender congruent intrusions in the “congruent lure” condition were compared. A higher proportion of induced congruent than induced incongruent intrusions was expected. This should result in a significant interaction between lure condition and response type, which was in fact observed (min. $LRCSS1 = 10.28$, $df = 1$, $p < .01$; min. $LRCSS2 = 9.28$, $df = 1$, $p < .01$).

The analyses excluding six items with substantial frequency differences between gender congruent and gender incongruent lure words revealed the same significant two-way interaction (min. $LRCSS1 = 4.54$, $df = 1$, $p < .05$; min. $LRCSS2 = 4.74$, $df = 1$, $p < .05$). The three-way interaction including participants or items was significant in neither of these analyses (all $ps > .99$). As depicted in Figure 11, there were more correct responses in the incongruent (92.9%) than in the congruent lure condition (82.5%) and more induced intrusions in the congruent (10%) than in the incongruent lure condition (1.3%).

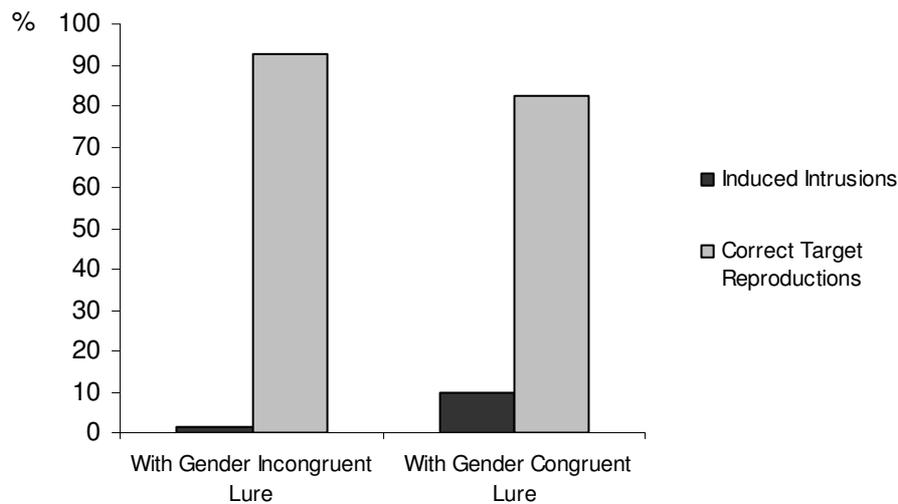


Figure 11: Percentages of induced lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 1.

4.3 Discussion

The observed pattern of data supports the assumption that grammatical gender information contributes to immediate sentence recall, at least with an auditory presentation modality. While a significant intrusion effect could be observed for gender congruent lure words, the proportions of correct responses and gender incongruent intrusions did not differ whether the incongruent lure word had been presented or not. The higher proportion of induced gender congruent than induced gender incongruent intrusions supports the assumption that gender information – as an example of morphosyntactic information – does contribute to short-term recall of sentences by reducing the lexical priming effect.

The higher proportion of induced gender congruent as compared to induced gender incongruent intrusions cannot be attributed to greater semantic similarity between the target word and its gender congruent competitor since both types of lure words were controlled for semantic similarity to the target word beforehand.

Nevertheless, the observed gender congruency effect might be due not to a contribution of morphosyntactic information, but simply to a contribution of phonological surface information. Phonological differences between target and gender incongruent noun phrase are larger than between target and gender congruent noun phrase because not only the nouns but also the *determiners* are different. The differ-

ence of phonological surface representations could prevent intrusions in the gender incongruent case if phonological information was available during recall.

Previous intrusion studies have demonstrated that phonological information is available during immediate recall if modality of sentence presentation is auditory (e.g., Rummer & Engelkamp, 2003). However, the involvement of phonological information depends on how sentences (or words) are presented.

Rummer and Engelkamp (2003) have compared recall of sentences presented in an auditory modality and via *rapid serial visual presentation* (RSVP). In RSVP, sentences are rapidly presented word-by-word at the center of the screen. In Rummer and Engelkamp's (2003) study, each word was displayed for 200 ms. They observed higher intrusion rates with RSVP than with auditory presentation. As suggested by Potter (1999), RSVP hampers phonological recoding, that is, RSVP should provide less phonological information as compared to auditory (or slow visual) presentation. With auditory presentation, phonological information is assumed to automatically enter phonological short-term memory, whereas visual verbal information has to be recoded by a rehearsal mechanism which can be conceptualized as a kind of inner speech linked to normal articulation (cf. e.g., Baddeley et al., 1975). The rehearsal rate is thus similar to an individual's normal speech rate. Hence, a presentation rate of 200 ms per word should be too fast to enable phonological recoding (for a more detailed discussion see Rummer & Engelkamp, 2001, p. 453f.). This is assumed to account for the different intrusion rates with sentences that have been presented in an auditory modality and with sentences that have been presented via RSVP.

In case the gender congruency effect for short-term memory observed in the first experiment reflected the use of phonological instead of morphosyntactic information, it should vary with the availability of phonological information and thus with modality of presentation. To test this alternative explanation, I have used RSVP instead of auditory presentation in a second experiment. As I assume the gender congruency effect to be independent of the availability of phonological information, the results obtained with auditory presentation are expected to occur even with RSVP, although weaker phonological traces are provided.

5. Experiment 2

This second experiment tested an alternative explanation for the gender congruency effect in sentence recall (i.e., for the observation of more induced gender congruent than induced gender incongruent intrusions) that has occurred in Experiment 1. It might be that the effect goes back to the contribution of phonological rather than morphosyntactic information. To exclude this, the experiment has been replicated with RSVP, a presentation mode that has been demonstrated to provide less phonological information than auditory presentation.

As I assume retention of gender information to be the basis of the findings from Experiment 1, the hypotheses for this experiment are the same as for the first one. An intrusion effect for gender congruent lure words is expected, while this is not sure for gender incongruent lure words. If the gender congruency effect at least in part goes back to a phonological representation of the target determiner, the intrusion effect for gender incongruent lure words might become significant with RSVP. However, the critical assumption is that the proportion of induced gender congruent intrusions should be higher than the proportion of induced gender incongruent intrusions, as was the case in the first experiment.

Experiment 2 was based on the same one-factorial design as Experiment 1. Within participants was manipulated whether the distractor list consisted only of unrelated words (*without lure*), whether a *gender incongruent lure word* was included in the list, or whether a *gender congruent lure word* was presented.

5.1 Method

Participants. 24 students of Saarland University, all of them native speakers of German, participated in this experiment. They either were awarded course credit or were paid for participation. None of them had participated in any of the other experiments.

Materials. The same materials as in Experiment 1 were used.

Procedure. The experimental procedure was equivalent to the first experiment. The only difference concerned the modality of presentation. To diminish the amount of phonological information available, sentences, word lists, and probe words were all presented via RSVP. Each trial was started by pressing the space bar of the computer keyboard. Three asterisks then appeared for 300 ms at the center of the screen followed by a blank screen for 350 ms. Then the word list was presented via RSVP at a

presentation rate of 250 ms per word. After an interval of 250 ms, in which a row of percentage signs was displayed, the sentence was presented via RSVP at a presentation rate of 200 ms per word. Subsequent to sentence presentation, participants had to recall the sentence orally as accurately as possible. Finally, the probe word was presented in capitals on the screen for 500 ms and participants had to decide via keystroke whether it had been part of the word list or not. The trial structure and the dependent variables were equivalent to Experiment 1. As in the first experiment, means and standard deviations for participants' overall accuracy of sentence recall are listed in Appendix B.

5.2 Results

Table 3 reports absolute frequencies and percentages of the four different response types (i.e., correct target reproductions, gender incongruent lure intrusions, gender congruent lure intrusions, and target omissions) per lure condition (i.e., list without lure, list with gender incongruent lure, and list with gender congruent lure). As with Experiment 1, an intrusion effect was observed only for gender congruent lure words. In addition, there was a higher proportion of induced gender congruent than gender incongruent intrusions whereas in the "without lure" condition there was obviously no difference between incongruent (0.4%) and congruent lure intrusions (2.1%).

Table 3: Absolute frequencies and percentages (in brackets) of gender congruent and gender incongruent lure intrusions, correctly reproduced target words, and target omissions as a function of lure condition in Experiment 2.

	Without Lure	With Gender Congruent Lure	With Gender Incongruent Lure
Gender Congruent Intrusions	5 (2.1%)	24 (10%)	9 (3.8%)
Gender Incongruent Intrusions	1 (0.4%)	1 (0.4%)	5 (2.1%)
Correct Target Reproductions	214 (89.2%)	195 (81.3%)	211 (87.9%)
Target Omissions	20 (8.3%)	20 (8.3%)	15 (6.3%)

The statistical analyses follow the same rationale as those of Experiment 1. That is, separate analyses testing for an intrusion effect for gender congruent lure words and for gender incongruent lure words, respectively, and an analysis that compares induced gender incongruent and gender congruent intrusions will be reported. As in Experiment 1, there was no difference between the three lure conditions with respect to target omissions (8.3% vs. 6.3% vs. 8.3%, Friedman test: $p > .3$).

(1) *Intrusion effect for gender congruent lure words.* The first analyses compared the correct responses and the gender congruent intrusions in those cases in which no lure word had been presented and in those in which the gender congruent lure word occurred on the list. As in Experiment 1, the expected two-way interaction between lure condition and response type reached significance (min. $LRCS1=7.80$, $df=1$, $p < .01$; min. $LRCS2=7.05$, $df=1$, $p < .01$) and this interaction was also significant when six items (critical with respect to frequency differences between congruent and incongruent lure words) were excluded from analysis (min. $LRCS1=6.40$, $df=1$, $p < .05$; min. $LRCS2=6.40$, $df=1$, $p < .05$). In addition, the effect did not interact with either the

participant or item factor (all $ps > .97$). Gender congruent intrusions were more frequent if the lure word had been presented on the distractor list (10%) than if no lure word had been presented (2.1%). The target word was reproduced more often in the “without lure” condition (89.2%) than in the “gender congruent lure” condition (81.3%). Figure 12 illustrates the intrusion effect for gender congruent lure words.

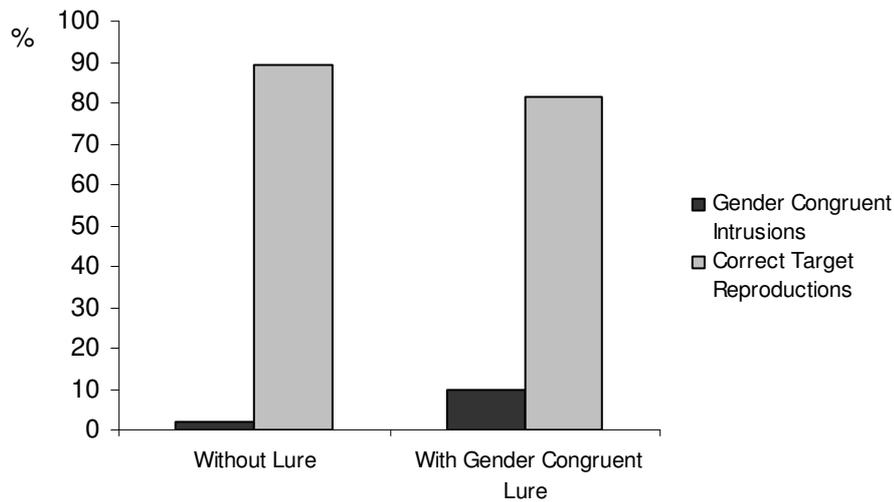


Figure 12: Percentages of gender congruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 2.

(2) *Intrusion effect for gender incongruent lure words.* Log-linear analyses including incongruent lure words did not reveal a significant two-way interaction between lure condition and response type (all $LRCs < 1$, $df = 1$, all $ps > .4$), which suggests that there was no significant difference between spontaneous (0.4%) and induced intrusions for gender incongruent lure words (2.1%). The results are depicted in Figure 13.

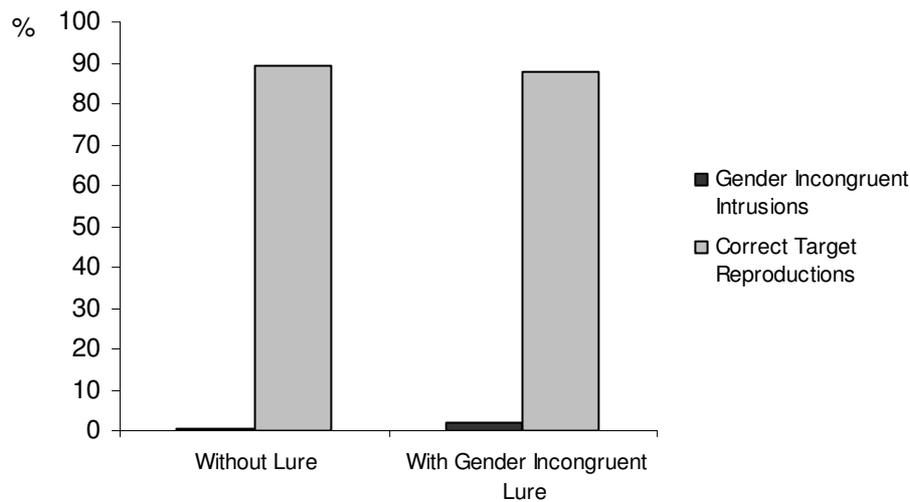


Figure 13: Percentages of gender incongruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 2.

(3) *Induced gender incongruent vs. induced gender congruent intrusions.* Crucially, there was a higher proportion of induced congruent than induced incongruent intrusions, which was suggested by a significant interaction between lure condition and response type (min. $LRC S1=7.54$, $df=1$, $p<.01$; min. $LRC S2=6.40$, $df=1$, $p<.01$). Again, the analyses with six items excluded revealed a significant two-way interaction (min. $LRC S1=3.68$, $df=1$, $p=.05$; min. $LRC S2=4.02$, $df=1$, $p<.05$) and the three-way interaction between lure condition, response type, and participants (or items) was not significant (all $ps>.99$). There were more correct responses in the incongruent (87.9%) than in the congruent lure condition (81.3%) and more induced intrusions in the congruent (10%) than in the incongruent lure condition (2.1%). Presentation of a gender congruent lure word led to more intrusions than presentation of a gender incongruent lure word, which is illustrated in Figure 14.

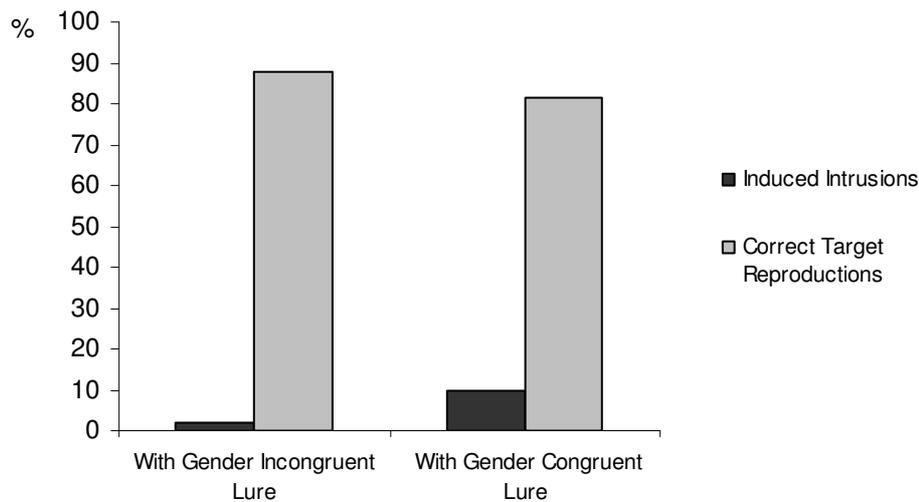


Figure 14: Percentages of induced lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 2.

5.3 Discussion

As with auditory presentation, the RSVP data support the assumption that gender information can contribute to verbatim recall of sentences. An intrusion effect, that is, an increase in intrusions as a consequence of lexical priming did not occur if there was a gender mismatch between target and lure word, but could be found if target and lure word were of the same grammatical gender.

Experiment 2 was conducted to test whether the gender congruency effect observed in Experiment 1 was due to a contribution of phonological information or, as predicted, due to a contribution of morphosyntactic (i.e., grammatical gender) information. In the present experiment the same critical results as in Experiment 1 were obtained. The (unmodified) occurrence of a gender congruency effect even under a condition in which substantially less phonological information is provided gives reason to assume that it is in fact the contribution of gender information that underlies the observed effect.

Nevertheless, the assumption that RSVP hampers phonological recoding which was made by Potter (1999, see above) was questioned by Coltheart (1999a, b). In various experiments, she demonstrated a phonological similarity effect not only with slow visual presentation but also with RSVP. As phonological similarity effects are usually interpreted in terms of an involvement of a phonological short-term store

(e.g., Baddeley, 1986; Conrad, 1964), this suggests that some phonological information is also provided by RSVP. Thus, it cannot yet be totally excluded that a phonological representation of the target determiner is responsible for the reduced amount of induced gender incongruent as compared to induced gender congruent intrusions.

It is well known that surface representations – in particular phonological information – decay rather quickly (e.g., Baddeley, 1986; cf. also Caplan, 1972; Jarvella, 1971; Sachs, 1967). According to Baddeley's phonological loop model, phonological information which is not subject to rehearsal decays after about two seconds. That is, phonological information should be available only under conditions in which recall immediately follows presentation.

Thus, the next experiment wants to make use of this fast decay rate. In this respect it is interesting that in most of Potter and Lombardi's (1990) original experiments, sentence recall was briefly delayed by placing presentation of the word list and the probe decision task in between sentence presentation and recall. Rummer and Engelkamp (2001, 2003) have argued that, if this procedure is applied, phonological information that might have been supplied by the time of sentence *presentation*, would no longer be available by the time of sentence *recall*.

To summarize, there are two ways to reduce the degree to which phonological information can contribute to recall. First, less phonological information can be provided by using a rapid visual presentation mode (i.e., RSVP). This procedure has been applied in Experiment 2. Second, recall can be delayed so that it can be assumed that the rapidly decaying phonological information is no longer available. This implies that even if some amount of phonological information has been provided with RSVP, it should have decayed if the sentence is not recalled until the intervening task (i.e., processing of the word list and probe decision) has been performed, that is, in *delayed recall*.

Therefore, in order to further exclude a phonological origin of the effect that has been attributed to morphosyntactic information, the experimental procedure was changed to a delayed recall situation in the third experiment. This procedure has already been applied in some of Potter and Lombardi's (1990) original experiments. It minimizes the access to phonological information during presentation (via RSVP) and during recall (via postponing recall). In addition, a possible gender congruency

effect under these conditions would suggest that morphosyntactic information is available for a longer period of time than is phonological information.

The above-mentioned studies that have compared immediate and delayed sentence recall in the intrusion paradigm have demonstrated an overall higher proportion of lure intrusions in delayed than in immediate recall. This is attributed to the decay of surface information that is not available to suppress intrusions any longer. Thus, one can expect more intrusions than in the previous experiments. Nevertheless, the intrusion effect should still be attenuated or eliminated for gender incongruent lure words, but should be observable for gender congruent lure words.

6. Experiment 3

In the third experiment, the influence of phonological information is further reduced by a delay between sentence presentation and sentence recall, which emanates through a change in the sequential order of sentence and list presentation. I assume that the gender congruency effect observed in the immediate recall experiments (Experiments 1 & 2) is caused by retention of grammatical gender information and not by phonological information. So there should still be a gender congruency effect in delayed recall of sentences presented via RSVP.

In general, a higher proportion of intrusion errors should be observed because the availability of phonological information is again reduced. A phonological representation of the target word is thus less able to prevent the lure words from intruding. This increase in intrusions might lead to a significant intrusion effect for gender incongruent lure words in this experiment, which was not the case in immediate recall. However, the critical comparison concerns induced gender congruent and induced gender incongruent lure intrusions. More induced congruent intrusions would once more constitute a gender congruency effect.

Experiment 3 was based on the same one-factorial design as the previous experiments. The only difference concerned the sequential order of sentence presentation, list presentation, and sentence recall. As in Experiments 1 and 2, it was manipulated whether the distractor list consisted only of unrelated words (*without lure*), whether a *gender incongruent lure word* was included in the list, or whether a *gender congruent lure word* was presented.

6.1 Method

Participants. 30 students participated in the experiment. All participants were native speakers of German and none of them had participated in one of the other experiments reported here. They were either paid for participation or received course credit.

Materials. The same materials as in Experiments 1 and 2 were used in this experiment.

Procedure. As in the second experiment, sentences, word lists, and probe words were presented via RSVP. The sequential order differed from the previous experiments. In the first two experiments an immediate recall procedure was used, whereas Experi-

ment 3 applied a delayed recall procedure. Each trial was started by pressing the space bar of the computer keyboard. Three asterisks appeared for 300 ms in the center of the screen followed by a blank screen for 350 ms. The sentence was presented via RSVP at a presentation rate of 200 ms per word. After an interval of 250 ms, in which a row of percentage signs was displayed, the word list was presented with 250 ms per word. For 517 ms the row of percentage signs appeared and afterwards the probe word was presented in capitals for 500 ms. After the probe judgment via key-stroke, participants had to recall the sentence orally as accurately as possible. The subsequent example (13) illustrates the trial structure with a gender incongruent lure-target relation (English glosses are given in brackets).

- (13) Sentence Presentation: Sie betraten voller Freude den *Rasen*_[masc] um Gänseblümchen und Klee zu pflücken
(They happily entered the *lawn*_[masc] in order to pick daisies and clover)
- List (with Incongruent Lure Word): Gruß Fähre Ball *Wiese*_[fem] Sprache
(accuracy ferry ball *meadow*_[fem] language)
- Probe Word: FÄHRE (FERRY)
- Sentence Recall (with Intrusion): Sie betraten voller Freude die *Wiese*_[fem] um Gänseblümchen und Klee zu pflücken
(They happily entered the *meadow*_[fem] in order to pick daisies and clover)

With the exception of the sequential order within the trials, Experiment 3 was identical to Experiment 2. With respect to accuracy of sentence recall, the results for Experiment 3 are given in Appendix B.

6.2 Results

Absolute frequencies and percentages of the four different response types (i.e., correct target reproductions, gender incongruent intrusions, gender congruent intrusions, and target omissions) per lure condition (i.e., list without lure, list with gender incon-

gruent lure, and list with gender congruent lure) are listed in Table 4. This time, in contrast to Experiments 1 and 2, a tendency towards an intrusion effect for gender incongruent lure words could be observed in addition to an intrusion effect for gender congruent lure words. Nevertheless, the proportion of induced gender congruent intrusions was substantially higher than the proportion of induced gender incongruent intrusions. Again, in the without lure condition, there was no significant difference between incongruent (4.7%) and congruent lure intrusions (6.7%). In addition, the three lure conditions did not influence the amount of target omissions (13.3% without lure vs. 14.7% with incongruent lure vs. 10.7% with congruent lure, Friedman test: $p > .6$)

Table 4: Absolute frequencies and percentages (in brackets) of gender congruent and gender incongruent lure intrusions, correctly reproduced target words, and target omissions as a function of lure condition in Experiment 3.

	Without Lure	With Gender Congruent Lure	With Gender Incongruent Lure
Gender Congruent Intrusions	20 (6.7%)	54 (18.0%)	22 (7.3%)
Gender Incongruent Intrusions	14 (4.7%)	10 (3.3%)	27 (9.0%)
Correct Target Reproductions	226 (75.3%)	204 (68.0%)	207 (69.0%)
Target Omissions	40 (13.3%)	32 (10.7%)	44 (14.7%)

(1) *Intrusion effect for gender congruent lure words.* The correct responses and the gender congruent intrusions in those cases in which no lure word had been presented and in those in which the gender congruent lure word occurred on the list were included in the first analysis. As in the first experiments, there was a significant two-

way interaction between lure condition and response type (min. $LRC S1=12.12$, $df=1$, $p<.001$; min. $LRC S2=12.12$, $df=1$, $p<.001$).

This interaction was also significant in the analyses in which the items with the six most frequent gender congruent lure words were excluded (min. $LRC S1=8.21$, $df=1$, $p<.01$; min. $LRC S2=8.86$, $df=1$, $p<.01$). In neither of these analyses an interaction with the participant or item factor was observed (all $ps>.4$). As expected, the gender congruent lure word intruded sentence recall more often if it was presented on the distractor list (18.0%) than if not (6.7%). In contrast, correct target reproductions were more frequent in the “without lure” condition (75.3%) than in the “gender congruent lure” condition (68.0%). Once more, an intrusion effect for gender congruent lure words was revealed by the analyses (see Figure 15).

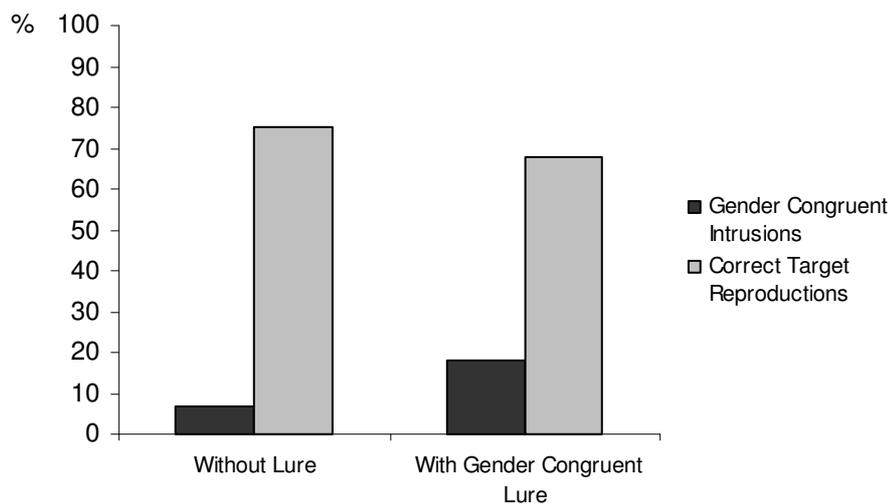


Figure 15: Percentages of gender congruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 3.

(2) *Intrusion effect for gender incongruent lure words.* In contrast to the first experiments, log-linear analyses including incongruent lure words revealed at least a tendency towards a two-way interaction between lure condition and response type (min. $LRC S1=3.11$, $df=1$, $p<.08$; min. $LRC S2=2.72$, $df=1$, $p<.1$; in the analysis with six items excluded: min. $LRC S1=2.42$, $df=1$, $p<.14$; min. $LRC S2=2.20$, $df=1$, $p<.12$). There was a numerical difference between spontaneous (4.7%) and induced intrusions (9.0%) for gender incongruent lure words. In addition, the proportion of correct

responses was reduced in the “incongruent lure” condition (69.0%) as compared to the “without lure” condition (75.3%). The results are depicted in Figure 16.

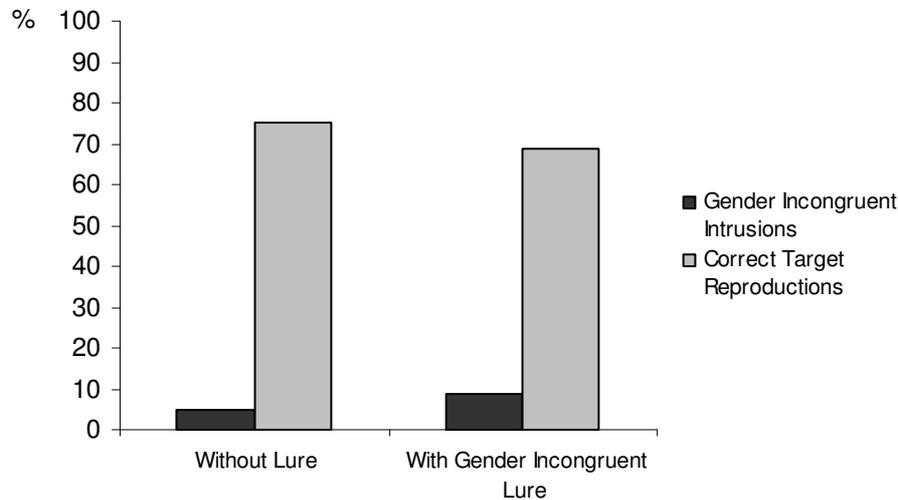


Figure 16: Percentages of gender incongruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 3.

(3) *Induced gender incongruent vs. induced gender congruent intrusions.* Particularly as more induced gender incongruent intrusions (than in Experiments 1 and 2) were observed in this experiment, the critical question is whether the proportion of induced congruent intrusions is still higher than the proportion of induced incongruent intrusions which would indicate a larger intrusion effect for gender congruent lure words. This was confirmed by a significant interaction between lure condition and response type (min. $LRCS1=5.61$, $df=1$, $p<.05$; min. $LRCS2=5.61$, $df=1$, $p<.05$).

However, in contrast to the first experiments this interaction did not reach significance when the six items with the largest frequency differences between congruent and incongruent lure words were excluded from the analyses (min. $LRCS1=1.54$, $df=1$, $p>.2$; min. $LRCS2=1.80$, $df=1$, $p>.18$).

As in the two immediate recall experiments, participants substituted the target word more often with a previously presented congruent lure word (18.0%) than with a previously presented incongruent lure word (9.0%). In this experiment, this did not result in a different proportion of correctly reproduced target words (which was the case in Experiments 1 and 2, see Figure 17). In the “incongruent lure” condition, 69.0% of the responses were correct, in the “congruent lure” condition the target was

correctly reproduced in 68.0% of the cases. The three-way interaction between lure condition, response type and participants or items, respectively, was not significant (all $ps > .18$).

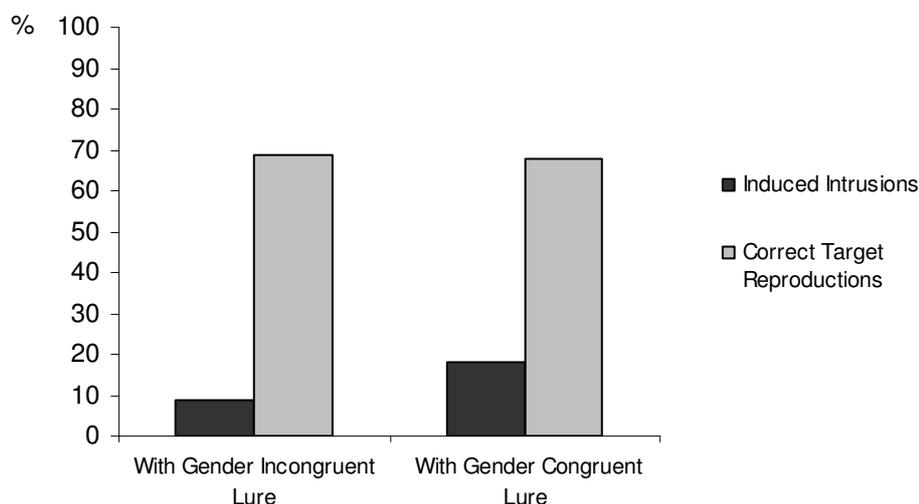


Figure 17: Percentages of induced lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 3.

6.3 Discussion

As in the previous experiments, the critical interaction between lure condition and response type when correct responses and induced intrusions were analysed revealed a higher proportion of induced gender congruent than of induced gender incongruent intrusions. This demonstrates a gender congruency effect even when the availability of phonological information is minimized in presentation *and* recall. Thus, also the results of this third experiment suggest that morphosyntactic information can contribute to short-term recall of sentences.

However, the findings of this experiment differ from the first two experiments in two respects: first, in addition to the intrusion effect for gender congruent lure words, in delayed recall following RSVP there was also a tendency towards an intrusion effect for gender incongruent lure words. Second, the analyses with six items excluded did not reveal a significant gender congruency effect.

Whereas the gender mismatch between target and lure word had made the intrusion effect disappear completely in immediate recall both after auditory presenta-

tion and after RSVP, when recall was delayed, there were again more induced gender congruent than gender incongruent intrusions, but numerically there were also more induced than spontaneous gender incongruent lure intrusions. It can be assumed that in a delayed recall condition there is hardly any phonological information available which could have been used to suppress retrieval of conceptually more adequate lexical competitors so that the probability for intrusion errors is higher overall. This is supported by the finding that the overall proportion of intrusions was higher in this delayed recall experiment than with immediate recall. The higher probability of intrusions in general then could have increased the probability of gender incongruent intrusions, too. Nevertheless, even under these conditions the proportions of induced gender congruent and incongruent intrusions still differed.

The second difference to Experiments 1 and 2 concerns the analyses in which the items with the six most frequent gender congruent lure words were excluded (and statistical power was thus reduced). In this experiment, only one of the critical analyses revealed a significant effect. Although there was an intrusion effect for gender congruent lure words, the difference between induced congruent and incongruent intrusions was no longer significant. However, the analyses indicating an intrusion effect for gender incongruent intrusions did not ‘come closer’ to significance in this analysis than in the analysis in which all items were included. Thus, the evidence for a gender congruency effect independent of frequency differences is weaker in this experiment than in the previous ones but the presence of an intrusion effect for gender congruent lure words in the absence of an effect for gender incongruent lure words in these additional analyses still indicates a recall influence of grammatical gender information.

The manipulations of modality of presentation and of the delay between sentence presentation and sentence recall aimed to reduce the influence of phonological information. Yet the above-mentioned study from Park and Martin (2002, see also Subchapter 1.4.3) gives reason to assume that phonological information can still influence recall even with delayed recall of sentences presented via RSVP. The authors observed a higher proportion of lure intrusions from words that were not only semantically but also phonologically related to the target word than from words that bore only a semantic relationship to the target word. This difference was smaller when recall was delayed but could still be observed. Thus it cannot yet be completely ex-

cluded that a phonological representation of the target determiner is responsible for the reduced amount of induced gender incongruent as compared to gender congruent intrusions. A contribution of the target determiner's *lexical* representation cannot be ruled out so far either.

However, certain characteristics of German determiners provide the opportunity to investigate cases in which target and gender incongruent lure words do not differ with respect to determiner form. For two of the three German gender types, the determiners differ or are the same depending on case. In the genitive and dative case, masculine and neuter nouns have the same definite determiner forms ("des" and "dem", respectively), whereas in the nominative and the accusative case, definite determiner forms are gender specific ("der" vs. "das" and "den" vs. „das“, respectively). While „der Hund/the dog“_[masc, nominative] and „das Kind/the child“_[neut, nominative] are both gender and determiner incongruent, „dem Hund“_[masc, dative] and „dem Kind“_[neut, dative] are gender incongruent but determiner congruent.

This allows for a manipulation of (un-)ambiguous morphophonological specification for gender in order to determine more specifically the locus of the gender congruency effect in short-term sentence recall. Hence, in the fourth experiment an additional factor was introduced, namely, whether or not the target determiner was adequate for both types of lure words. In other words, it was manipulated whether the morphophonological gender specification of the target determiner was unambiguous (as in nominative and accusative case) or ambiguous (as in genitive and dative case).

7. Experiment 4

This experiment aimed at clarifying whether the gender congruency effect is really a gender congruency effect or whether it only demonstrates that the target determiner can be remembered even though the target noun cannot. A lexical representation of the target determiner could then help to reject gender incongruent intrusions but not gender congruent intrusions. If the latter, the gender congruency effect should be influenced by the morphophonological specification of the determiner. Determiner specification has been varied in this experiment (i.e., it has been varied whether or not the target determiner matched the determiner of the gender incongruent lure).

If this determiner-based explanation was right, there should again be more induced gender congruent than gender incongruent intrusions in the condition in which the determiner is unambiguously specified for gender (as a replication of the first experiments). Yet an intrusion effect of comparable size should be observable for gender congruent and gender incongruent lure words when the determiner is the same for the target word, the gender congruent lure word, *and* the gender incongruent lure word.

In contrast, if what has been termed a gender congruency effect is indeed a gender congruency effect, it should still be observed under these conditions. That is, an intrusion effect for gender congruent but not (necessarily) for gender incongruent lure words should occur and, critically, more induced gender congruent than induced gender incongruent intrusions should be produced.

The experiment was based on a two-factorial design, with both *lure condition* (without lure, with gender congruent lure, and with gender incongruent lure) and *morphophonological specification for gender* (determiner ambiguously specified, determiner unambiguously specified) varied within participants.

7.1 Method

Participants. 42 students of Saarland University participated in this experiment. All participants were native speakers of German and none of them had participated in one of the other experiments. They participated for course credit or were paid.

Materials. A set of 30 sentences and word lists was constructed. For each of the 30 sentences two versions existed. In one version, the target determiner was the same as the determiner of the respective gender incongruent lure word. In the other version,

the gender incongruent lure word had a different determiner than the target. Only masculine and neuter target and lure words were used. For these gender types genitive and dative determiners have the same form and nominative and accusative determiners have different forms. For most of the sentences the determiner manipulation was operationalized as follows: in one version of the sentences, the target word was in nominative or accusative case, including a determiner uniquely specified for gender (see example (14a) below).

- (14a) Die Frau sah aus, als hätte sie *ein Phantom*_[neut, accusative] gesehen, und lief davon.

Literal translation: The woman looked as if had she a *spirit*_[neut, accusative] seen and ran away.

Gloss: The woman looked as if she had seen a *spirit* and ran away.

Gender congruent lure word: *Gespenst*_[neut] (phantom)

Gender incongruent lure word: *Geist*_[masc] (ghost)

In this case, a different determiner has to be used if a gender incongruent intrusion occurs. An intrusion without a change of the determiner would either lead to an agreement error or, as in (14a), to a change in thematic roles. In the other version, either the verb or a preposition was changed so that the target word was in genitive or dative case. This implies that both the gender congruent and the gender incongruent lure word had the same determiner as the target word (see example (14b) below). Here, just the same determiner can be used even in the case of a gender incongruent intrusion.

- (14b) Die Frau sah aus, als wäre sie *einem Phantom*_[neut, dative] begegnet, und lief davon.

Literal translation: The woman looked as if had she a *spirit*_[neut, dative] come across and ran away.

Gloss: The woman looked as if she had come across a *spirit* and ran away.

Gender congruent lure word: *Gespenst*_[neut] (phantom)

Gender incongruent lure word: *Geist*_[masc] (ghost)

Each participant was presented with 15 sentences of each version and across participants each sentence appeared equally often in each version. In addition, as in the first experiments, for one third of the sentences, the accompanying word list included a control word, for one third a gender congruent lure word, and for one third a gender incongruent lure word.

Pre-tests. The lure words in this set of materials were also controlled for lexical-semantic and phonological factors. The results of the pre-tests will be briefly described in the following paragraphs (means and standard deviations for each pre-test are given in Table 5 below).

Table 5: Means and standard deviations (in brackets) for the target and lure words in the pre-tests for Experiment 4.

	Target Word	Gender Congruent Lure Word	Gender Incongruent Lure Word
Semantic Similarity (1-10)	--	5.9 (<i>SD</i> =1.7)	5.8 (<i>SD</i> =1.5)
Cloze Probabilities (%)	1.7 (<i>SD</i> =3.6)	28.0 (<i>SD</i> =23.8)	29.8 (<i>SD</i> =19.0)
Word Frequency (per million)	378 (<i>SD</i> =955)	455 (<i>SD</i> =549)	377 (<i>SD</i> =549)
Word Length (Characters)	7.0 (<i>SD</i> =1.9)	6.1 (<i>SD</i> =2.1)	6.2 (<i>SD</i> =2.2)
Word Length (Syllables)	2.1 (<i>SD</i> =0.7)	2.0 (<i>SD</i> =0.7)	1.9 (<i>SD</i> =0.9)
Shared Phonemes (Final Syllable)	--	0.7 (<i>SD</i> =0.8)	0.5 (<i>SD</i> =0.6)

20 participants completed the sentences with the target determiners and nouns omitted. On average, the sentences were completed with the incongruent lure words in 29.8 % of the cases, with the congruent lure words in 28.0%, and with the target words in 1.7% of the cases. 20 more participants rated the semantic similarity be-

tween lure and target words (on a scale ranging from 1 to 10). On average, the gender incongruent lure words were rated equally similar to the target words (5.8) as the gender congruent lure words (5.9).

In addition, there was no difference in word length between the gender congruent (6.1 characters and 2.0 syllables) and the gender incongruent lure words (6.2 characters and 1.9 syllables). Neither did the gender congruent and the gender incongruent lure words differ in rhyme overlap with the target word. On average, the final syllables of target and gender congruent lure words overlapped in 0.5 phonemes in the same position and in 0.1 more phonemes in different positions. Target and gender incongruent lure words' final syllables shared 0.3 phonemes in the same position and 0.2 more in different positions.

Although with respect to word frequency there was a numerical difference between the gender congruent (455 per million) and the gender incongruent lure words (377 per million), this difference was not significant ($t(29) = -.65$; $p > .5$). A complete list of the sentences in both versions and of the respective lure words can be found in Appendix A.

Procedure. The experimental procedure was the same as in Experiment 2 and the same dependent variables were analyzed. Again, means and standard deviations of participants' overall accuracy in sentence recall are given in Appendix B.

7.2 Results

Absolute frequencies and percentages of the four different response types (i.e., correct target reproductions, gender incongruent lure intrusions, gender congruent lure intrusions, and target omissions) per lure condition (i.e., list without lure, list with gender incongruent lure, and list with gender congruent lure) and morphophonological specification condition (i.e., determiner ambiguously specified for gender, determiner unambiguously specified for gender) are listed in Table 6. As in the previous experiments, an intrusion effect was observed only for the gender congruent lure words and the proportion of induced gender congruent intrusions was substantially higher than the proportion of induced gender incongruent intrusions. Critically, both effects were independent of the specification of the determiner. Again, in the without lure condition, there was no significant difference between incongruent (1.9%) and congruent lure intrusions (3.6%).

Table 6: Absolute frequencies and percentages (in brackets) of gender congruent and gender incongruent lure intrusions, correctly reproduced target words, and target omissions as a function of lure condition and morphophonological specification for gender in Experiment 4.

	Target Determiner Unambiguously Specified for Gender			Target Determiner Ambiguously Specified for Gender		
	Without Lure	With Gender Congruent Lure	With Gender Incongruent Lure	Without Lure	With Gender Congruent Lure	With Gender Incongruent Lure
Gender Congruent Intrusions	9 (4.3%)	16 (7.6%)	7 (3.3%)	6 (2.9%)	25 (11.9%)	6 (2.9%)
Gender Incongruent Intrusions	2 (1.0%)	2 (1.0%)	3 (1.4%)	6 (2.9%)	3 (1.4%)	10 (4.8%)
Correct Target Reproductions	192 (91.4%)	179 (85.2%)	187 (89.1%)	181 (86.2%)	170 (81.0%)	178 (84.8%)
Target Omissions	7 (3.3%)	13 (6.2%)	13 (6.2%)	17 (8.1%)	12 (5.7%)	16 (7.6%)

The statistical analyses follow the same rationale as in the previous experiments. I will first report the analyses indicating whether there is an intrusion effect for gender congruent and for gender incongruent lure words and will then proceed to the analysis that compares induced gender congruent and induced gender incongruent intrusions. The amount of target omissions was neither influenced by the lure condition (5.7% without lure vs. 6.0% with gender congruent lure vs. 6.9% with gender incongruent lure, Friedman test: $p > .5$) nor by the morphophonological specification of the target determiner (7.1% with ambiguous determiner vs. 5.2% with unambiguous determiner, Friedman test: $p > .4$).

(1) *Intrusion effect for gender congruent lure words.* The correct responses and the gender congruent intrusions in those cases in which no lure word had been presented and in those in which the gender congruent lure word occurred on the list were included in the first separate analysis. As in the previous experiments, there was a significant two-way interaction between lure condition and response type (min. $LRCs_1 = 5.57$, $df = 1$, $p < .05$; min. $LRCs_2 = 6.61$, $df = 1$, $p < .05$), which was not influenced by the participant or item factor (all $ps > .99$). As expected, the gender congruent lure word intruded sentence recall more often if it had been presented on the distractor list (9.8%) than if not (3.6%). In contrast, correct target reproductions were more frequent in the “without lure” condition (88.8%) than in the “gender congruent lure” condition (83.1%). Once more, an intrusion effect for gender congruent lure words was revealed by the analyses (see Figure 18 below). Moreover, whether there was a mismatch between the determiner for the target and the determiner for the gender incongruent lure word did not affect this intrusion effect (all $LRCs < 1$, $df = 1$, all $ps > .7$).

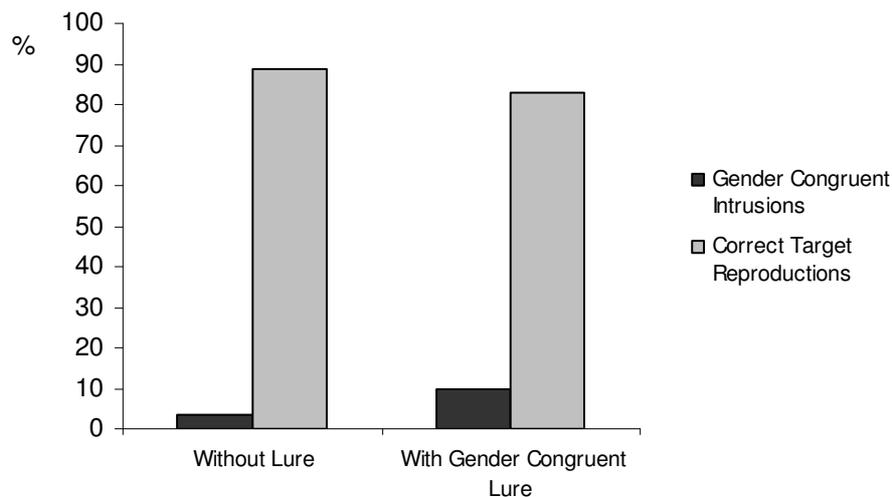


Figure 18: Percentages of gender congruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 4.

(2) *Intrusion effect for gender incongruent lure words.* In the log-linear analyses including incongruent lure words neither the two-way interaction between lure condition and response type nor the three-way interaction between lure condition, response type, and morphophonological specification reached significance (all $LRCs < 1$, $df = 1$, all $ps > .5$). That is, a significant difference between spontaneous (1.9%) and induced intrusions (3.1%) for gender incongruent lure words was observable in neither specification condition. In addition, the proportion of correct responses did not differ in the “without lure” condition (88.8%) and in the “incongruent lure” condition (86.9%). These results are illustrated in Figure 19 below.

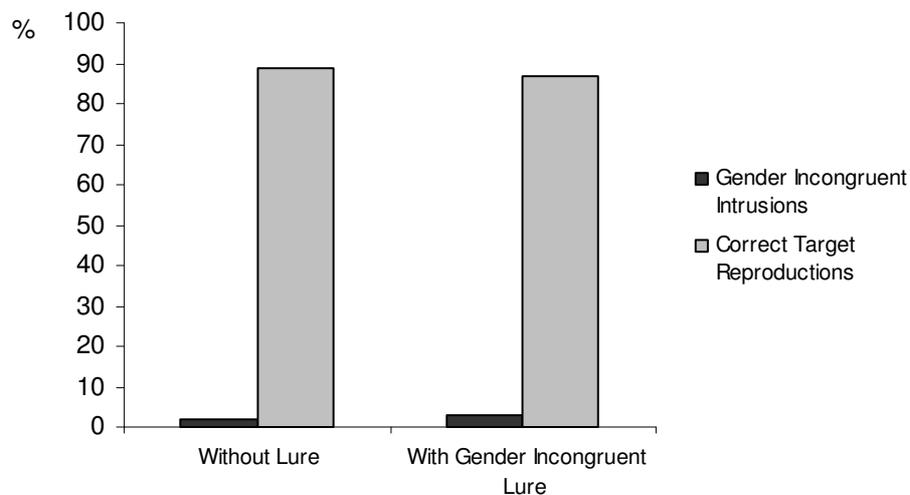


Figure 19: Percentages of gender incongruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 4.

(3) *Induced gender incongruent vs. induced gender congruent intrusions.* The critical question is whether the proportion of induced congruent intrusions is higher than the proportion of induced incongruent intrusions, which would indicate a larger intrusion effect for gender congruent lure words. This was confirmed by a significant interaction between lure condition and response type (min. $LRCs1=5.48$, $df=1$, $p<.05$; min. $LRCs2=6.55$, $df=1$, $p<.05$) in the absence of a three-way interaction between lure condition, response type, and participants or items (all $ps>.99$) and of a three-way interaction between lure condition, response type, and morphophonological specification for gender (all $LRCs <1$; all $ps>.7$). As in the first experiments, participants substituted the target word more often with a previously presented congruent lure word (9.8%) than with a previously presented incongruent lure word (3.1%). Correct reproductions of the target word were slightly more frequent in the “incongruent lure” condition (86.9%) than in the “congruent lure” condition (83.1%). As illustrated in Figure 20, this analysis again revealed a gender congruency effect.

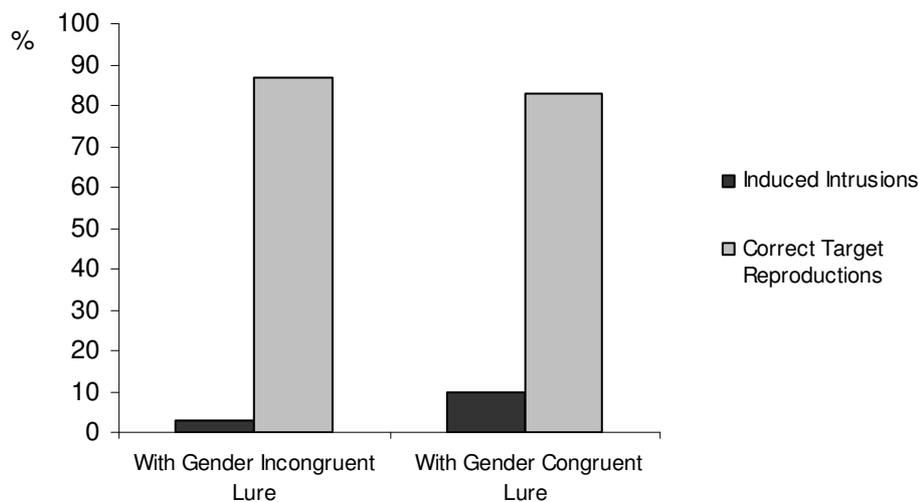


Figure 20: Percentages of induced lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 4.

In addition to the gender congruency effect, this analysis revealed a weak tendency toward a two-way interaction between specification and response type (min. $LRCs1=2.56$, $df=1$, $p<.11$; min. $LRCs2=3.03$, $df=1$, $p<.09$). Surprisingly, both gender congruent and gender incongruent intrusions were (numerically) more frequent when the determiner was compatible with both lure words than if it did not match the gender incongruent one.

7.3 Discussion

This experiment provided the opportunity to rule out the possibility that the gender congruency effect goes back solely to an influence of the target determiner. So it was the most critical test for the assumption that morphosyntactic information is represented in short-term memory. In fact, the results of the other immediate recall experiments (Experiments 1 & 2) could be replicated. There was an intrusion effect for the gender congruent lure words but not for the gender incongruent lure words and there were more induced gender congruent than induced gender incongruent intrusions. This gender congruency effect was independent of whether or not there was a determiner mismatch between the target word and the gender incongruent lure word, which clearly demonstrates that it is not the retention of the target determiner's surface form that causes the reduced error rate for gender incongruent lure words. Instead, there has to be some other mechanism that leads to more gender congruent than gender incongruent induced semantic errors. What this mechanism might be will be discussed below (cf. General Discussion).

It should be mentioned that, in addition to the gender congruency effect, the analyses including induced gender congruent and gender incongruent intrusions and correct target reproductions revealed a slight tendency toward an interaction between morphophonological specification and response type. Induced intrusions of both gender types were more frequent when the determiner was compatible with the gender congruent and the gender incongruent lure word. Such an increase in intrusions could have been expected for the gender incongruent lure words but is rather counterintuitive for the gender congruent lure words. Since I do not have a convincing explanation for this result at the moment and since the interaction effect was not significant, I will leave it at this for now.

Experiments 1 to 4 have investigated the contribution of grammatical gender information to short-term recall of sentences. In the following subchapter, I will jointly discuss the results of these experiments before proceeding to the final experiment. This experiment aims at extending the gender congruency effect to recall of more complex materials.

7.4 Joint Discussion of Experiments 1-4

The four experiments reported so far have each revealed a gender congruency effect in short-term recall of sentences – irrespective of modality of presentation, of the delay between sentence presentation and sentence recall, and of whether or not there was a determiner mismatch between the target word and the gender incongruent lure word. In each of the experiments, the presentation of a lure word that was semantically related to a target word in the sentence and had the same grammatical gender as the target word led to a substantial amount of lure intrusions. These lure intrusions were more frequent as compared to a condition in which the list consisted solely of words that bore no semantic relationship to the target word. In contrast, if a word was presented that was semantically related to the target but had a different grammatical gender this word replaced the target no more than when it had not been presented. This was also reflected in a higher proportion of induced gender congruent as compared to induced gender incongruent lure intrusions, the so-termed gender congruency effect. The same pattern of results was observed across all four experiments. Although the overall proportion of intrusion errors increased from immediate sentence recall (Experiments 1, 2, and 4) to delayed sentence recall (Experiment 3), this did not influence the relative proportion of gender congruent and gender incongruent intrusion errors. This indicates that grammatical gender (representing morphosyntactic information) plays a prominent role in short-term verbatim recall of syntactically structured materials. The gender congruency effect neither depends on the availability of phonological representations of the critical region of the sentence (i.e., target word and target determiner) nor on the contribution of a lexical representation of the target determiner.

Thus, the present findings extend previous evidence for a close link between representations relevant in language processing and representations involved in verbal short-term memory tasks, beyond representations that can be investigated on the single word level. Several studies have found that phonological information and lexical-semantic information contribute to short-term retention of isolated words in lists and embedded words in sentence contexts (e.g., Conrad, 1964; Watkins, 1977). The present study demonstrates the influence of morphosyntactic information.

A representation of the grammatical gender of the nouns presented in a to-be-recalled sentence is used to prevent recall intrusions of semantically similar but mor-

phosyntactically different nouns. Morphosyntactic information can eliminate the impact of lexical priming but nevertheless influences sentence recall only in collaboration with semantic information. This is reflected in the absence of intrusions from gender congruent distractor words (on the list) other than the semantically related lure word.

The assumption that activated parts of the language processing system are the basis of verbal short-term memory (which underlies all language-based models) questions certain aspects of “separatist” models of verbal working memory. For instance, it questions the existence of a separate surface representation stored in short-term memory and a conceptual or gist representation stored in long-term memory. Such a view cannot explain the mutual influence and interdependence of representations traditionally associated with long-term memory such as conceptual information or the mental lexicon (e.g., Potter & Lombardi, 1990) and of representations classified as surface information such as phonological information (Park & Martin, 2002; Rummer & Engelkamp, 2001, 2003) or grammatical gender information.

The present results underline the importance of models that emphasize the close connection between language processing and short-term memory for language and they underline the importance of providing theories and data on more complex stimuli. The investigation of morphosyntactic information particularly requires the application of syntactically structured materials. Serial recall of word lists has been successfully applied to investigate the recall contribution of phonological and lexical-semantic features. However, morphosyntactic information is different from phonological and lexical-semantic information. A main difference concerns its relational aspects. The impact of grammatical gender becomes obvious in syntactic contexts. It relates determiners and nouns, adjectives and nouns, or nouns and anaphoric pronouns. An influence of grammatical gender should thus indeed be most prominent in recall of sentences.

What is more, the account of verbal short-term memory that I have promoted here (see also Subchapter 1.4.3 above) does not only suggest that activated parts of the language processing system are critical for verbal short-term memory. It also includes assumptions about what kind of representations stay activated and how long the storage of certain representations should last. As stated above, the general idea is that what is retained in short-term memory is what is or might be necessary for fur-

ther language processing and that what might be necessary for further language processing stays activated in short-term memory. In other words, the requirements of language processing determine the contents and characteristics of verbal short-term memory.

What is meant by the requirements of language processing? For instance, in normal conversations one needs a representation of a sentence's message in order to successfully integrate new incoming information. Since in many cases the content of several sentences has to be available to completely understand a new sentence, such representations need to be maintained relatively long. On the contrary, it is normally not essential for understanding subsequent parts of discourse to retain information about constituent order of a processed sentence (whether it says "the cat and the dog" or "the dog and the cat") or about lexical surface differences that result only in minimal meaning differences (whether it says "tired" or "sleepy").

As to grammatical gender information, which is examined here, I have already stated that it is often needed for determining the right antecedent for an anaphoric pronoun (or at least for narrowing down the search field of possible antecedents). However, this function of gender information is not limited to adjacent sentences. Although a sentence that intervenes between a noun and a pronoun that refers to the noun reduces the importance of retention of morphosyntactic information (e.g., Clark & Sengul, 1979), some information still has to be available. In such cases, one normally re-introduces the noun with a noun (phrase) rather than referring to it with a pronoun (e.g., Sanford & Garrod, 1981). This noun phrase just needs to match the antecedent with respect to content and does not need to agree with it morphosyntactically. Yet an anaphoric pronoun as in (15) can still be interpreted correctly, although interpretation may be harder. This can be done by means of access to morphosyntactic information (see also Garnham et al., 1995, Exp. 3).

(15) Ich beobachtete ein Huhn_[neut], das einen Wurm_[masc] jagte. Die Sonne_[fem] schien. Plötzlich kam *es*_[neut] auf mich zu.

Literal translation: I watched a chicken_[neut] that a worm_[masc] chased. The sun_[fem] shone. Suddenly came *it*_[neut] towards me.

Gloss: I was watching a chicken that was chasing a worm. The sun was shining. Suddenly *it* approached me.

Thus, grammatical gender information is still needed for anaphor resolution even when an intervening sentence separates the pronoun and its antecedent. This makes it plausible to assume that an influence of gender information on verbal retention should not be restricted to recall of single sentences but should also be observable if a critical target word is part of a sentence presented further back.

Such an assumption differs from other models of short-term memory for text – or rather for more than one sentence. As outlined in Subchapter 1.4.1, most models assume that only conceptual information is available beyond single sentences, whereas a representation of so-termed surface information is only available for the most recent sentence. Such a separation of long-term conceptual and short-term surface information has been supported by both recognition (Sachs, 1967, 1974; Caplan, 1972) and recall studies (Jarvella, 1971, 1979).

It was assumed that the ability to recognize or recall the exact wording of a sentence reflects the ability to retain surface information. The term surface information thus refers to representations that are quite heterogeneous in several of their processing characteristics. Most prominently, these are word choice and syntactic representations. Both representations of exact lexical choice and of surface syntactic structure of a given sentence (e.g., whether the sentence is in passive or in active voice or whether an object is expressed by a prepositional phrase or by a simple noun phrase) were shown to decay rapidly after a sentence had been processed (Caplan, 1972; Jarvella, 1971, 1979). However, Sachs (1974) observed different decay rates for lexical and for syntactic information. At a point of time when her participants still reliably detected syntactic changes from the original to the probe sentence, detection of lexical changes was already at chance level (cf. also Figure 6 above).

However, the underspecified treatment of surface information in these studies (with the exception of Sachs' experiments) may have obscured differences between certain types of surface information with respect to their availability and "sustainability" in short-term memory. Since grammatical gender information can clearly be classified as surface information, models in the tradition of Jarvella (1971) would not predict an influence of grammatical gender once processing of a sentence is completed.

By contrast, the account proposed here would still predict a contribution of gender information under conditions that eliminate the influence of other types of

surface information. However, this influence might be smaller than for single sentences or for the most recent sentence.

The final experiment addresses the question whether or not the gender congruency effect in short-term recall of syntactically structured materials is restricted to recall of single sentences and whether it is influenced by the state of processing of the critical information.

8. Experiment 5

In order to test whether the gender congruency effect extends to short-term recall of more than one sentence, the final experiment of this study applied the intrusion paradigm in its modified version (as used in Experiments 1 to 4) to a text recall task. More precisely, Jarvella's (1971, 1979) paradigmatic prose recall procedure was applied. As described in Subchapter 1.4.1, he had his participants recall prose passages. As soon as presentation of a text was interrupted, participants were supposed to recall as much of the previously perceived discourse as possible. Analysing recall accuracy for each clause and sentence, Jarvella found that the most recent sentence was recalled better than earlier sentences (taking accuracy of verbatim sentence recall as dependent variable). Critically, this was also the case for the penultimate clause. This clause could either be connected to the final clause and thus be part of the most recent sentence or connected to the previous clause and thus already be completely processed. Although its wording was the same in both conditions, recall for this critical clause was better in the former than in the latter case. A modification of this paradigm has been applied in Experiment 5.

This final experiment used short texts that ended in three-clause passages comparable to those in Jarvella's experiments. Each text existed in two versions: in one version the penultimate clause formed a sentence with the previous clause, in the other version with the final one. The wording of these three clauses was the same in both versions. In contrast to Jarvella's experiments, there was an additional manipulation in the materials: the critical middle clause contained a target word for which both a gender congruent and a gender incongruent lure word existed.

However, contrary to the sentence recall experiments these lure words were not presented as part of a separate word list. The use of text materials allowed for placing either lure word (or a control word in the "without lure" condition) within the text. In order to present the participants with texts that consisted of more than two sentences, the critical three-clause passage in which the target word was positioned (henceforth *target region*) had to be preceded by a few context sentences. The lure manipulation concerned the last sentence of this *context passage*. This sentence was always constructed such that it could contain the gender congruent lure word, the gender incongruent lure word, or an unrelated control word.

If necessity for further text processing does in fact guide activation in short-term memory, the processing characteristics of gender information should lead to a gender congruency effect. This effect should be comparable to the effects in the sentence recall experiments, even though more than one sentence has to be recalled and even though the critical target word is part of a sentence that has been followed by another one. That is, a gender congruency effect should be observed in this text recall experiment and it should occur no matter if the target word is part of the most recently processed sentence or part of a less recent sentence.

According to traditional theories of sentence memory (e.g., Jarvella, 1971), there should also be an increase in intrusion errors for the condition in which the target clause is part of the penultimate sentence as compared to the condition in which it belongs to the most recent sentence. Since on the basis of a conceptual representation it is hard to distinguish target and lure words (and lure words should even be favoured by the conceptual representation) and since lexical information should be less accessible for the former sentence, the lure word should substitute the target word more often in this condition. In general, this task is more difficult than the single-sentence recall task used in the previous experiments. Additionally, presentation of the lure lies further back than in these experiments. It is thus plausible to assume, that participants are able to correctly recall the target word less frequently and, in turn, omit the target more often.

In this experiment, the manipulations do not only concern the target and lure words. The state of processing of the penultimate clause (part of the most recent sentence vs. part of a less recent sentence) should also affect overall recall performance as was the case in Jarvella's experiments. Thus, the analyses for the subsequent experiment will not only focus on lure intrusions but will also address general recall accuracy. According to the assumption that a sentence's state of processing has an influence on how accurately it can be recalled, results similar to Jarvella's findings should be expected. That is, the second clause of the target region should be recalled more accurately when it is connected to the first clause than when it is connected to the final one. In addition to an effect of sentence boundaries, Jarvella (e.g., 1979) postulated an influence of clause boundaries (see also Caplan, 1972). That is, the final clause is assumed to be recalled more accurately than the two previous clauses. Obviously, the final clause is not only part of the most recent sentence but is also the

most recent clause. Therefore, the highest degree of recall accuracy in the present experiment can be expected for the final clause.

Experiment 5 was based on a two-factorial design with the two within-subject factors *lure condition* and *state of processing*. With respect to lure condition, it was manipulated whether the context sentences contained an unrelated control word (*without lure*), a *gender incongruent lure word*, or a *gender congruent lure word*. In addition, the middle one of the critical three clauses was either *connected to the final clause* or *connected to the previous clause (state of processing)*.

8.1 Method

Participants. 48 students of Saarland University participated in this experiment. They either received course credit or were paid. All participants were native speakers of German and none of them had participated in one of the previous experiments.

Materials. 30 short texts (consisting of four to seven sentences) were constructed such that each one ended in a three-clause structure. As in Jarvella's experiments, two versions of this *target region* existed: either the first and the second clause formed a sentence and the final clause stood alone or the first clause stood alone and the second and the third clause formed a sentence. Due to this, the critical middle clause was either part of the most recent sentence or part of a sentence back. In spite of this structure manipulation, the sentences' wording was the same in both versions except for obligatory word order differences in the final clause.

Each of the three-clause passages was preceded by a context passage consisting of two to five sentences. The length of these passages was varied so that participants could not guess which part of the text they had to recall, but each three-clause passage was presented with the same pretext in each condition. Similar to the first four experiments, the middle clause contained a not optimally fitting target word for which two better fitting lure words existed. Again, one of these lure words was gender congruent to the target word and one was gender incongruent. In contrast to the first four experiments, the lure words were not presented in a list preceding the critical sentence. Instead, items were constructed such that either one of the lure words or a control word appeared in the sentence preceding the critical region (see example (16) below: (a) indicates the condition in which the target clause belongs to the previous clause and (b) indicates the condition in which it belongs to the subsequent

clause; the critical clause, lure and control words in the context passage are italicized). Subsequent to the German original, literal translations and glosses are given. All three-clause passages and the respective lure words are listed in Appendix A.

(16) *Context passage*: Denise hatte sich für den Samstag fest vorgenommen, einen Stadtbummel zu machen. Am frühen Morgen schon ließ sie sich mit ihrem *Auto*_(gender congruent lure word, neut)/*Wagen*_(gender incongruent lure word, masc)/*Hund*_(control word) abholen und nahe der Fußgängerzone absetzen.

Target region:

- (a) Der Parkplatz füllte sich langsam,
während der Chauffeur das *Fahrzeug*_(target word, neut) bewachte.
Denise ging zum Einkaufen in die Innenstadt.
- (b) Der Parkplatz füllte sich langsam.
Während der Chauffeur das *Fahrzeug*_(target word, neut) bewachte,
ging Denise zum Einkaufen in die Innenstadt.

Literal translation:

Context passage: Denise had herself for the Saturday definitely planned a stroll through the town to make. Ate early morning already let she herself with her *automobile*_(gender congruent lure word, neut)/*car*_(gender incongruent lure word, masc)/*dog*_(control word) pick up and near the pedestrian mall drop.

Target region:

- (a) The car park filled itself slowly
while the chauffeur the *vehicle*_(target word, neut) watched.
Denise went for shopping into the city.
- (b) The car park filled itself slowly.
While the chauffeur the *vehicle*_(target word, neut) watched,
Denise went for shopping into the city.

Glosses:

Context passage: For Saturday, Denise had planned to make a stroll through the town. Early in the morning she was picked up with her *automobile*/*car*/*dog* and was dropped near the pedestrian mall.

Target region:

- (a) The car park filled up slowly
while the chauffeur was watching the vehicle.
 Denise went downtown for shopping.
- (b) The car park filled up slowly.
While the chauffeur was watching the vehicle,
 Denise went downtown for shopping.

Pre-tests. As with the other materials, it had to be ensured that grammatical gender was the only critical difference between the two types of lure words. The results of the pre-tests are listed in Table 7.

Table 7: Means and standard deviations (in brackets) for the target and lure words in the pre-tests for Experiment 5.

	Target Word	Gender Congruent Lure Word	Gender Incongruent Lure Word
Sem. Similarity (1-10)	--	5.9 (<i>SD</i> =1.7)	5.7 (<i>SD</i> =1.5)
Cloze Probabilities (%): target in penulti- mate sentence	0.8 (<i>SD</i> =1.9)	22.0 (<i>SD</i> =22.7)	30.3 (<i>SD</i> =21.7)
Cloze Probabilities (%): target in ultimate sentence	1.0 (<i>SD</i> =2.4)	22.0 (<i>SD</i> =20.7)	28.3 (<i>SD</i> =23.5)
Word Frequency (per million)	561 (<i>SD</i> =1296)	431 (<i>SD</i> =559)	389 (<i>SD</i> =440)
Word Length (Characters)	6.5 (<i>SD</i> =2.3)	6.5 (<i>SD</i> =3.1)	6.5 (<i>SD</i> =2.1)
Word Length (Syllables)	2.1 (<i>SD</i> =0.9)	2.0 (<i>SD</i> =0.8)	2.1 (<i>SD</i> =0.8)
Shared Phonemes (Final Syllable)	--	0.6 (<i>SD</i> =0.8)	0.6 (<i>SD</i> =0.7)

In the first pre-test, 20 native speakers of German (none of whom participated in one of the experiments) rated the semantic similarity between lure and target words on a scale ranging from 1 (no semantic similarity) to 10 (high semantic similarity). On average, the gender incongruent lure words were rated equally similar to the target words (5.7) as the gender congruent lure words (5.9).

In addition, both types of lure words had to fit the sentence content equally well, which was measured by a sentence completion task. 40 more participants completed the sentences with the target determiners and nouns omitted. The entire three-clause passages in each version were completed by 20 participants each. On average, the sentences were completed with the incongruent lure words in 30.3% of the cases when the middle clause was connected to the first clause and in 28.3% of the cases when it was connected to the final clause. In both versions, the congruent lure words were filled in in 22% of the cases, respectively. That is, cloze probability differed neither between the two versions nor between the two types of lure words (with a slight advantage for the incongruent lure words). The target word was used in 0.8 % of the cases when the critical clause belonged to the previous one and in 1.0% of the cases when it belonged to the last one.

Furthermore, the two lure types matched with respect to their word length in characters and in syllables. On average, the incongruent lure words consisted of 6.5 characters and 2.1 syllables and the congruent lure words consisted of 6.5 characters and 2.0 syllables.

Mean frequency of the gender incongruent lure words was 389 per million and for the gender congruent lure words 431 per million. With respect to phonemic overlap there was no difference between gender congruent and gender incongruent lure words either. On average, the final syllables of the gender congruent lure words shared 0.5 phonemes in the same position and 0.2 more phonemes overall with the target word; for gender incongruent lure words these were 0.4 and 0.1 phonemes, respectively.

Procedure. For each participant, a third of the texts included a gender congruent lure word, one third included a gender incongruent lure word, and one third included a control word that was not related to the target word. Across participants, each text appeared equally often with a gender congruent lure word, a gender incongruent lure

word, and a control word. In addition, each participant was presented with 15 texts in which the critical clause was connected to its preceding clause and with 15 texts in which it was connected to its subsequent clause. Text assignment to one of the state of processing conditions was counterbalanced across participants.

Each trial was started by pressing the space bar of the computer keyboard. Three asterisks then appeared at the center of the screen for 300 ms followed by a blank screen for 350 ms. Then the text was presented auditorily in normal speech rate via loudspeakers. After a 500 ms break, in which a row of percentage signs was displayed, an auditory recall cue indicated the starting point for text recall. The recall cue was always the beginning (up to the first noun) of the critical three-clause region. After the cue, participants had to recall the last two sentences of the text (from the cue on) as accurately as possible.

Before the experiment started, three practice trials were presented to the participants. The experiment lasted about half an hour. Participants' utterances were recorded and transcribed. The number (and the proportion) of correct reproductions of the target word, of gender incongruent lure intrusions, and of gender congruent lure intrusions served as dependent variables. In addition, the number (and the proportion) of target omissions was recorded.

Recall accuracy of the second and of the third clause will be reported as a further dependent measure. The first clause of the target region has not been included in this analysis because the beginning of this clause was used as an auditory recall cue. I will report two measures of overall recall accuracy: first, the proportion of correctly reproduced words per clause, irrespective of order and with the target noun phrases excluded from analysis; second, the proportion of clauses that have been recalled perfectly. Correct reproduction of the target word is irrelevant for this measure. In other words, a clause is regarded as perfectly recalled even if the target word has been substituted by one of the lure words.

8.2 Results

The results section for Experiment 5 is divided into two main parts: first, the results with respect to lure intrusions will be reported, the first part of which is comparable to the previous experiments. Additionally, separate analyses for the conditions in which the target clause was part of the most recently encountered sentence and for

the conditions in which it was part of a less recent sentence will be reported. The second part of this results section addresses the question of general recall accuracy and includes analyses on recall performance for each clause as a function of its state of processing.

8.2.1 Lure Intrusions

In contrast to the sentence recall experiments, there was a large proportion of cases in which neither the target word nor one of the lure words was produced (i.e., target omissions). An analysis of target omissions revealed a main effect for the state of processing of the target clause (min. $LRC\text{S}1=5.57$, $df=1$, $p<.05$; min. $LRC\text{S}2=6.38$, $df=1$, $p<.05$): there were more target omissions when the target clause belonged to the penultimate sentence (25.1%) than when it belonged to the ultimate sentence (18.2%). In addition, there was a main effect for item (min. $LRC\text{S}2=138.64$, $df=26$, $p<.001$), which implies that certain items caused several target omissions, while others caused hardly any. In addition, the omissions of some items were often due to a complete failure to recall the target clause or even the entire target region. In other words, the items differed extremely in their degree of difficulty. One possibility to reduce the large differences in the item set is to exclude the most difficult items from analyses. A descriptive analysis of target omissions revealed that for six of the 30 items the proportion of omission errors was more than one standard deviation above the mean across items (20.8%, $SD=19.1$). For these six items, the target was omitted by more than 40% of the participants. All relevant analyses concern the relative proportion of intrusion errors and correct reproductions of the target word. A large proportion of target omissions reduces both the number of correct target reproductions and lure intrusions *unsystematically*. In other words, a proportion of target omissions of 40% and more induces a large amount of noise. This is particularly the case as each participant is presented with only five items per condition. This noise may hinder the detection of the critical interactions. I will therefore report analyses from which these six items have been excluded but will mark the cases in which the results differ from an analysis including all 30 items. Detailed results of these complete analyses can be found in Appendix C. The analyses on target omissions with 24 of the 30 items still revealed a main effect of item ($LRC\text{S}2=45.19$, $df=20$, $p<.01$). None of the other main effects or interactions reached significance.

Table 8 reports absolute frequencies and percentages of the four different response types (i.e., correct target reproductions, gender incongruent intrusions, gender congruent intrusions, and target omissions) per lure condition (i.e., without lure, with gender incongruent lure, and with gender congruent lure) and state of processing condition (i.e., target belongs to the penultimate sentence and target belongs to the ultimate sentence). In contrast to the single sentence recall experiments, there were intrusion effects (i.e., more lure intrusions if the context passage included a lure word as compared to cases in which it included an unrelated control word instead) for both gender congruent and for gender incongruent lure words. However, as expected, the proportion of induced intrusions was higher for gender congruent than for gender incongruent lure words, whereas there were hardly any spontaneous intrusions in both the gender congruent (7.8%) and the gender incongruent cases (3.6%). All these interactions were independent of the state of processing of the critical clause that included the target word.

Table 8: Absolute frequencies and percentages (in brackets) of gender congruent and gender incongruent lure intrusions, correctly reproduced target words, and target omissions as a function of lure condition and state of processing in Experiment 5.

	Target Word Part of the Penultimate Sentence			Target Word Part of the Ultimate Sentence		
	Without Lure	With Gender Congruent Lure	With Gender Incongruent Lure	Without Lure	With Gender Congruent Lure	With Gender Incongruent Lure
Gender Congruent Intrusions	17 (8.9%)	49 (25.5%)	11 (5.7%)	13 (6.8%)	43 (22.4%)	12 (6.3%)
Gender Incongruent Intrusions	7 (3.7%)	2 (1.0%)	29 (15.1%)	7 (3.7%)	3 (1.6%)	15 (7.8%)
Correct Target Reproductions	127 (66.2%)	118 (61.5%)	127 (66.2%)	148 (77.1%)	125 (65.1%)	141 (73.4%)
Target Omissions	41 (21.4%)	23 (12.0%)	25 (13.0%)	24 (12.5%)	21 (10.9%)	24 (12.5%)

In the following, I will first report analyses in which both “state of processing” conditions are included. As in the previous experiments, the first of these analyses addresses the question of an intrusion effect for gender congruent lure words (1a). The second type of analysis concerns an intrusion effect for gender incongruent lure words (1b) and the last of these analyses assesses differences between induced gender congruent and induced gender incongruent intrusions (1c), analogously to the procedure in the first experiments.

Subsequently, the findings for each “state of processing” condition will be reported separately. That is, I will report analyses addressing an intrusion effect for gender congruent lure words (2a), an intrusion effect for gender incongruent lure words (2b), and a gender congruency effect for the condition in which the target clause is part of the penultimate sentence (2c). In addition, the equivalent analyses (paragraphs (3a) to (3c)) for the condition in which the target clause is part of the ultimate sentence will be reported.

(1) General Analyses

(1a) Intrusion effect for gender congruent lure words. To analyze whether there was a significant intrusion effect for gender congruent lure words, the correct responses and the gender congruent intrusions were compared when either no lure word or the gender congruent lure word was presented in the text. An intrusion effect would be reflected in a two-way interaction between lure condition (without lure vs. with congruent lure) and response type (correct target reproduction vs. congruent lure intrusion). Log-linear analyses revealed the expected two-way interaction (min. $LRCs1=18.46$, $df=1$, $p<.001$; min. $LRCs2=23.73$, $df=1$, $p<.001$). Participants produced more gender congruent intrusions if the lure word had been presented in the context passage (24.0%) than if not (7.8%). In contrast, they reproduced the target more often in the “without lure” condition (71.6%) than in the “gender congruent lure” condition (63.3%), which indicates an intrusion effect for gender congruent lure words (see Figure 21). In addition, the three-way interaction between lure condition, response type, and state of processing was not significant (all $LRCs<1$, $df=1$, all $ps>.5$). This indicates that the intrusion effect was not modulated by the syntactic status of the target clause. Although the three-way interaction with participants was not significant in this experiment either (both $ps>.97$), inter-item variation was

stronger than in the sentence recall experiments and the interaction with items approached significance (both $ps > .1$).

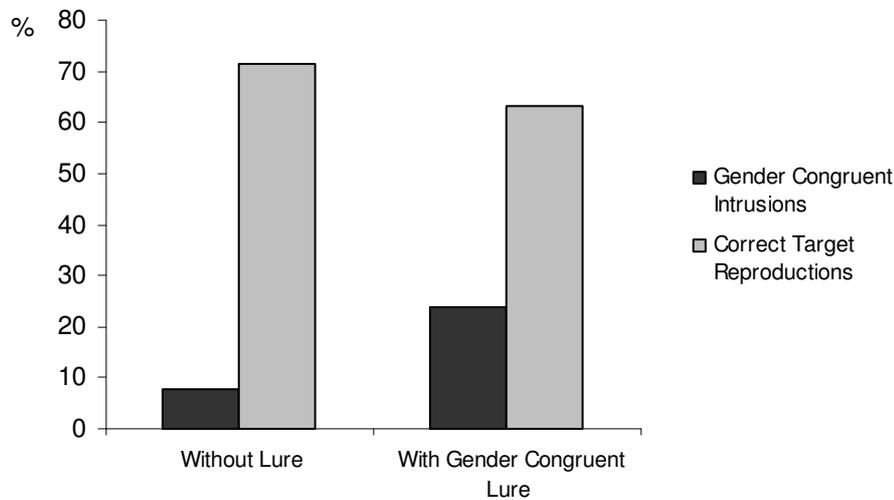


Figure 21: Percentages of gender congruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 5.

(1b) *Intrusion effect for gender incongruent lure words.* There was an intrusion effect for the equivalent analyses regarding gender incongruent intrusions (including the “without lure” and the “gender incongruent” condition and the correct responses and the gender incongruent intrusions). Log-linear analyses revealed a significant two-way interaction between lure condition and response type (min. $LRCs1=5.28$, $df=1$, $p<.05$; min. $LRCs2=7.36$, $df=1$, $p<.01$). There were more gender incongruent intrusions when the incongruent lure word was part of the context passage (11.5%) than when an unrelated word was presented instead (3.6%). Correct target reproductions were more frequent in the “without lure” (71.6%) than in the “incongruent lure” condition (69.8%). This interaction is depicted in Figure 22. It was neither modulated by the state of processing of the target clause (all $LRCs < 1.3$, $df=1$, all $ps > .2$) nor by the participant or item factor (all $ps > .98$).

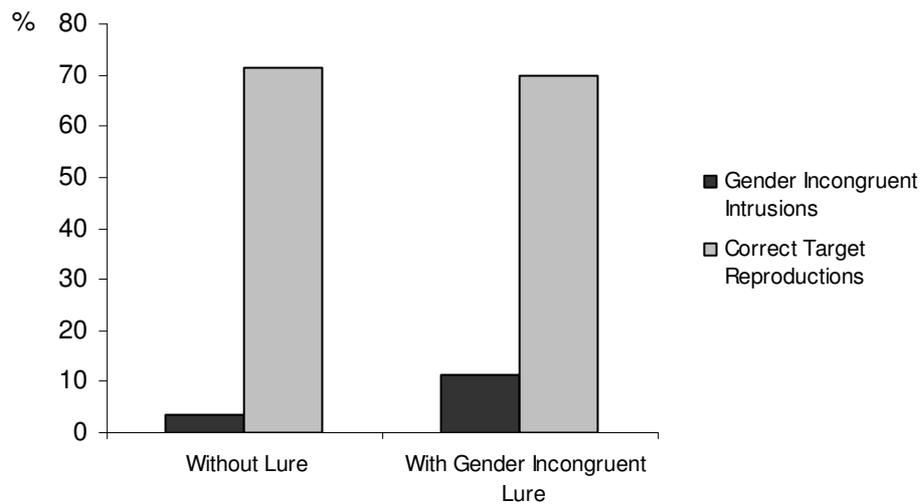


Figure 22: Percentages of gender incongruent lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 5.

(1c) *Induced gender incongruent vs. induced gender congruent intrusions.* This analysis allowed for a comparison of the impact that presentation of the respective lure word had on gender congruent and gender incongruent intrusions. Here, correct responses and gender incongruent intrusions in the “incongruent lure” condition as well as correct responses and gender congruent intrusions in the “congruent lure” condition were included in the analyses. A higher proportion of induced congruent than induced incongruent intrusions was expected and could in fact be observed. This was reflected in a significant interaction between lure condition and response type (min. $LRCS1=9.36$, $df=1$, $p<.01$; min. $LRCS2=13.07$, $df=1$, $p<.001$, see Figure 23). Correct responses were produced more often in the incongruent (69.8%) than in the congruent lure condition (63.3%) and induced intrusions more often in the congruent (24.0%) than in the incongruent lure condition (11.5%). Again, the three-way interaction between lure condition, response type, and state of processing did not reach significance (both $LRCS1s<1$, $df=1$, both $ps>.4$; both $LRCS2s<1.5$, $df=1$, both $ps>.2$). The three-way interaction between lure condition, response type, and participants was not significant either (both $ps>.99$) but the analysis with items suggests some inter-item variation (both $ps>.16$).

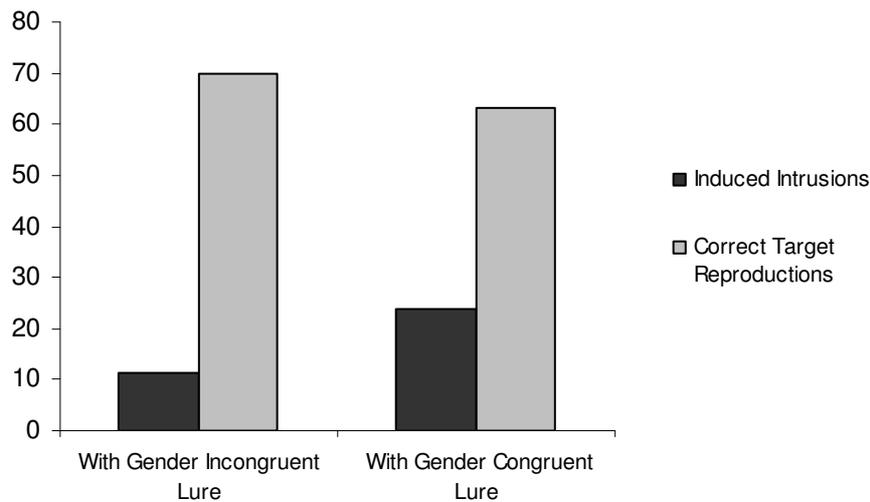


Figure 23: Percentages of induced lure intrusions and correct reproductions of the target word as a function of lure condition in Experiment 5.

(2) *Separate analyses for the conditions in which the target word is part of the ultimate sentence*

(2a) *Intrusion effect for gender congruent lure words.* These analyses were equivalent to those reported in paragraph (1a) with the exception of the state of processing of the target clause. Here, only the condition in which the critical clause was part of the most recent sentence was included. Again, an intrusion effect for gender congruent lure words was revealed, that is, there was a significant two-way interaction between lure condition and response type (min. $LRCS1=10.31$, $df=1$, $p<.01$; min. $LRCS2=13.18$, $df=1$, $p<.001$). Congruent lure intrusions were more frequent if the lure word was presented in the context passage (22.4%) than if not (6.8%). Correct reproduction of the target word succeeded in 65.1% of the cases in the congruent lure condition but in 77.1% of the cases in the without lure condition. Lure condition and response type did not interact with either the participant or item factor (all $ps>.79$).

(2b) *Intrusion effect for gender incongruent lure words.* These results differed from the more general analyses (1b). In the separate analyses including gender incongruent intrusions and correct responses in the “without lure” and in the “incongruent lure” condition, no intrusion effect for gender incongruent lure words could be observed when only the condition was analyzed in which the target word belonged to

the most recent sentence (min. $LRC S1=1.03$, $df=1$, $p>.3$; min. $LRC S2=1.32$, $df=1$, $p>.2$). There was only a numerical difference between spontaneous (3.6%) and induced gender incongruent intrusions (7.8%) and between correct responses in the “without lure” (77.1%) and in the “incongruent lure” condition (73.4%).

(2c) *Induced gender incongruent vs. induced gender congruent intrusions.* This analysis was similar to the more general analysis (1c). A gender congruency effect was indicated by a significant two-way interaction between lure condition and response type (min. $LRC S1=4.96$, $df=1$, $p<.05$; min. $LRC S2=10.25$, $df=1$, $p<.01$) in the absence of a three-way interaction with participants or items (all $ps>.98$). When the target word was part of the most recent sentence, induced gender congruent intrusions were produced more often (22.4%) than induced gender incongruent intrusions (7.8%). In addition, the target word was reproduced correctly more often when the incongruent lure word had been presented in the context passage (73.4%) than when the congruent lure word had been presented (65.1%).

(3) *Separate analyses for the conditions in which the target word is part of the penultimate sentence*

(3a) *Intrusion effect for gender congruent lure words.* When comparing congruent intrusions and correct target reproductions in the “without lure” and the “congruent lure” condition of those trials in which the target clause was part of the penultimate sentence, the two-way interaction between lure condition and response type was significant (min. $LRC S1=6.96$, $df=1$, $p<.01$; min. $LRC S2=10.46$, $df=1$, $p<.01$). The three-way interaction between lure condition, response type, and participants (both $ps>.99$) or items, respectively, was not significant (both $ps>.4$). Congruent intrusions were produced more often if the lure word had been presented in the context passage (25.5%) than if not (8.9%) whereas the target was correctly reproduced slightly more often in the “without lure” condition (66.2%) than in the “congruent lure” condition (61.5%).

(3b) *Intrusion effect for gender incongruent lure words.* In the equivalent analyses for gender incongruent intrusions, there was also an intrusion effect indicated by a two-way interaction between lure condition and response type (min. $LRC S1=4.58$,

df=1, $p < .05$; min. $LRCSS2=6.64$, df=1, $p < .01$). 3.7% intrusions in the “without lure” condition opposed 15.1% intrusions in the “incongruent lure” condition. The target was correctly reproduced in 66.2% of the cases when a control word was presented in the context passage and when the gender incongruent lure word was presented. The interaction was not influenced by the participant or the item factor (all $ps > .98$).

(3c) *Induced gender congruent vs. induced gender incongruent intrusions.* The critical analysis compared induced gender congruent and induced gender incongruent intrusions and correct target reproductions in the two conditions in which a lure word was presented in the text. This interaction was marginally significant when only the cases were included in which the target word was part of a sentence back (min. $LRCSS1=3.18$, df=1, $p < .08$; min. $LRCSS2=3.75$, df=1, $p = .05$). In this condition, participants produced 15.1% induced incongruent intrusions and 25.5% induced congruent intrusions. They correctly reproduced the target word in 66.2% of the cases and 61.5%, respectively. The three-way interaction with participants was not significant (both $ps < .99$) but the interaction with items reached significance (min. $LRCSS2=36.15$, df=1, $p < .05$). However, the two-way interaction between lure condition and response type did not reach significance when all 30 items were included in the analyses (min. $LRCSS1=2.00$, df=1, $p < .17$; min. $LRCSS2=2.33$, df=1, $p < .12$). The three-way interactions with participants or items, respectively, also revealed different results: both the interactions with participants (both $ps < .07$) and items (both $ps < .1$) approached significance.

8.2.2 Clause Recall

With respect to overall recall accuracy, I will first report analyses comparing recall accuracy for the second as compared to the third clause. Afterwards, separate comparisons of recall accuracy for the second clause and for the third clause as a function of the clauses' state of processing will be outlined. Two different measures will be reported: (1) the proportion of correctly reproduced words per clause (excluding the target word and determiner and irrespective of order); (2) the proportion of clauses that have been perfectly recalled (irrespective of whether or not the target has been reproduced). The first clause was never included in the analyses since the recall cue constituted a substantial part of that clause. For each comparison, Wilcoxon matched

pairs tests will be reported. This test is the non-parametric, distribution-free analogue to the *t*-test for related samples. The analyses also rely on 24 of 30 items. Unless indicated otherwise, the analyses with all 30 items reveal comparable effects. The results of these analyses are listed in Appendix C.

As expected, recall performance for the penultimate clause was better when it was part of the most recent sentence than when there was a sentence boundary before the final clause. However, contrary to Jarvella's results, there was no general advantage for the last clause. The penultimate clause was unexpectedly recalled more accurately than the ultimate clause, in both state of processing conditions. Figure 24 depicts the results with respect to the proportion of correctly reproduced words per clause; Figure 25 depicts the results with respect to the proportion of clauses that were recalled completely.

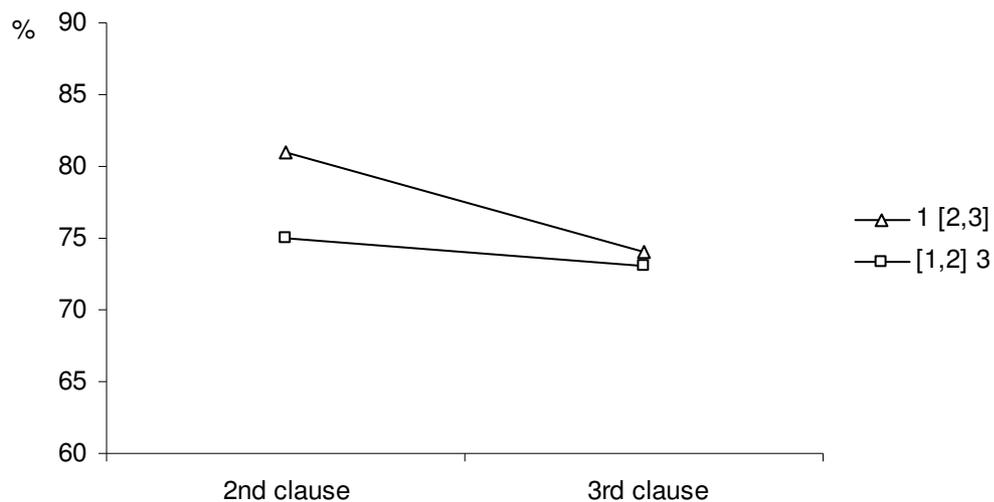


Figure 24: Proportion of correctly reproduced words per clause (excluding the target noun phrase) as a function of syntactic structure in Experiment 5.

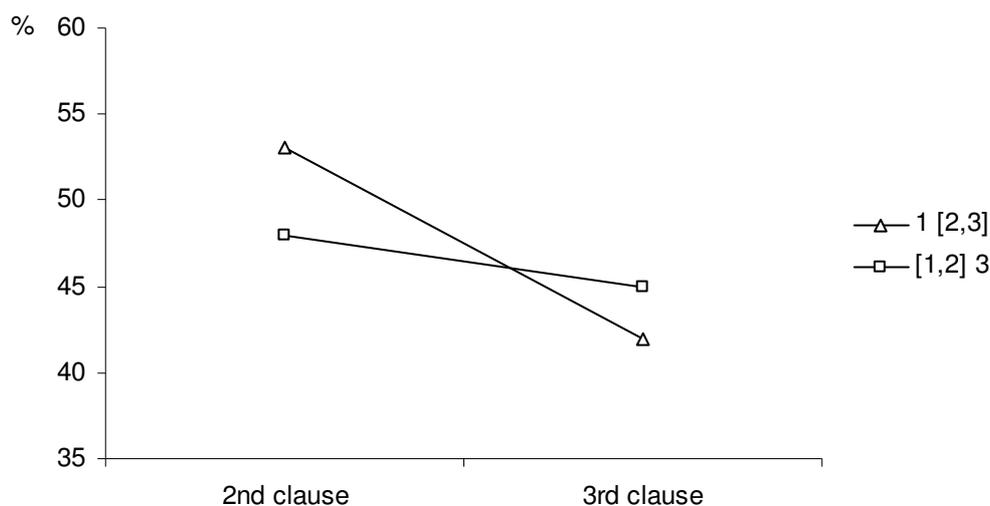


Figure 25: Proportion of perfectly reproduced clauses (irrespective of correct reproduction of the target word) as a function of syntactic structure in Experiment 5.

Recall accuracy: second vs. third clause. Unexpectedly, in both states of processing the second clause was recalled more accurately than the third clause. When the middle clause was part of the penultimate sentence, participants correctly reproduced 75% of its words and 73% of the words of the final clause. Though this difference was small, it was significant ($T=43157$, $z=5.6$, $p<.001$). When these two clauses formed a processing unit (i.e., a sentence), 81% of the words in the middle clause and 74% of the words in the final clause were correctly recalled ($T=50961$, $z=3.61$, $p<.001$).

When the proportion of clauses that were perfectly recalled was compared, this difference was not significant for the condition in which the middle clause belonged to the previous one ($T=13899$, $z=1.05$, $p>.2$). On average, 48% of the middle clauses and 45% of the final clauses could be recalled entirely when the middle and the final clause were separated by a sentence boundary. However, when the middle clause was part of the ultimate sentence, it was more often perfectly reproduced (53%) than the final clause (42%; $T=10604$, $z=3.58$, $p<.001$).

Recall accuracy as a function of state of processing. When comparing the proportion of words per clause that were recalled correctly, the state of processing influenced performance for the penultimate clause ($T=20800.5$, $z=3.83$, $p<.001$). Participants correctly recalled 75% of the words of the middle clause when it belonged to the

penultimate sentence and 81% when it belonged to the ultimate sentence. When the proportion of perfectly recalled clauses was analysed, this difference could be observed numerically (53% vs. 48%) but was not significant ($T=13695$, $z=1.54$, $p>.12$).

As expected, neither the proportion of correctly recalled words of the final clause ($T=34387.5$, $z=.15$, $p>.8$) nor the proportion of perfectly recalled clauses were influenced by the state of processing of the penultimate clause ($T=14022$, $z=.94$, $p>.3$). 73% of the words were recalled correctly if the ultimate clause constituted the entire sentence and 74% when the second and the third clause constituted the final sentence. In the condition in which the middle clause and the final clause were separated by a sentence boundary, the final clause was perfectly recalled in 45% of the cases as compared to 42% when they formed a sentence together. This indicates a numerical advantage for the condition in which the final clause was a sentence by itself, which even approached significance in the analysis with all 30 items included.

8.3 Discussion

The last experiment of the present study has investigated recall of multiple sentences within short prose passages. Continuously, participants had to update their memory representation of the text because they did not know when recall started and thus did not know beforehand which part of the text they would have to recall.

First of all, the results demonstrate that the lexical priming of the lure word does work when the lure is embedded in a running text and not only when it is presented on an extra distractor list. The intrusion effects for both lure types were significant in all but one of the analyses. It was the first time that the intrusion paradigm was applied to recall of longer discourse and the results are encouraging with respect to further applications.

More specifically, the experiment aimed at finding out whether the previously observed gender congruency effect in short-term memory extended to a text recall task and whether it was influenced by the state of processing of the clause the target word belonged to. This was operationalized by combining the intrusion paradigm with a prose recall procedure introduced by Jarvella (1971). It was assumed that grammatical gender information should be retained over sentence boundaries and that thus a gender congruency effect should also be observable in recall of short texts. In addition, the gender congruency effect should occur both when the target

word was part of the most recent sentence and when it was part of a sentence further back.

The findings do not support this absolutely but to a large extent. The finding of a significant gender congruency effect in the overall analysis, which was not modulated by the state of processing factor, supports the assumption that the recall contribution of grammatical gender information should be observable whether or not the target clause is part of the most recently encountered sentence. In addition, significant intrusion effects for gender congruent lure words and, contrary to the single-sentence recall experiments, for the gender incongruent lure words were observed. These interactions were not influenced by the state of processing of the target clause either.

A more direct way to test whether gender information is maintained until recall even with an additional sentence boundary between the critical noun phrase and recall was to analyse this (and the second) state of processing condition separately. The analyses including the subset of cases in which the target was in the most recent sentence also revealed a significant gender congruency effect. In other words, presentation of a gender congruent lure word induced more intrusions than presentation of a gender incongruent lure word. When only the subset of cases was analysed in which the target word was in the penultimate sentence (i.e., when there was an additional sentence boundary after the target clause), the critical interaction that implies a gender congruency effect was only marginally significant. This indicates that the results should be taken with a pinch of salt. However, in combination with the significant gender congruency effect in the overall analysis (i.e., when both the cases in which the target clause was connected to the previous clause and in which it was connected to the final clause were included), it can be concluded that to some degree grammatical gender information contributes to recall of syntactically structured materials even across sentence boundaries.

The text recall task that was applied in this experiment demanded a great deal of the participants. Since the paradigm had been newly introduced, the difficulty of the task in general and of certain items in particular could not sufficiently be assessed beforehand. However, even in its present form, conclusions can be drawn with respect to the role of grammatical gender information in short-term recall of short text passages. In sum, one can conclude that a joint representation of the meaning and the

grammatical gender of a noun (phrase) in a text is retained. The results of the separate analyses for the conditions in which the critical noun phrase was part of the most recent sentence and for the conditions in which it was part of the penultimate sentence indicate that the influence of the gender congruency decreases to some degree over sentence boundaries.

In a way, these results can be regarded as reflecting the actual use of grammatical gender information in language processing. I have argued that grammatical gender is still used in anaphor interpretation, as is the case in example (15) above. However, whether or not grammatical gender information suffices depends on the circumstances. In example (15), the intervening sentence does not introduce a new discourse topic. The more discourse topics intervene between antecedent and pronoun, the higher the probability that there is more than one possible antecedent that matches the pronoun in grammatical gender will be. And due to other constraints in anaphor resolution, the most *recent* gender congruent noun phrase would be the most probable antecedent for the pronoun. This suggests that although there is still a need for gender information to be accessible even after an intervening sentence, the influence clearly decreases (see also Clark & Sengul, 1979). Maybe a more pronounced gender congruency effect could be observed with materials that are designed so that few discourse objects are introduced in the final clause.

Such a “yes but less” influence of grammatical gender information after an intervening sentence is also reflected in the results of Garnham et al.’s (1995) final experiment (see also Chapter 2.1 above). In contrast to most studies that investigate the influence of gender information on pronoun interpretation, they dissociated the influence of conceptual and grammatical gender information. In a study language that provides grammatical gender information for genderless entities (here: French) they compared pronouns referring to people and pronouns referring to things. When there was an intervening sentence between the antecedent and the referring pronoun, participants were able to make use of the fact that the antecedent could uniquely be identified on the basis of a gender cue. Yet this advantage was more pronounced when the antecedent was a person (and thus marked with respect to grammatical *and* conceptual gender) than when it was a thing (and thus marked *solely* with respect to grammatical gender). When the pronoun was presented in a sentence that directly followed the sentence including the antecedent, the use of a gender cue was not in-

fluenced by the antecedent type (object vs. person; Garnham et al., 1995, Exp. 1 & 2). The present experiment was designed to exclude confounds between grammatical and conceptual gender and is thus comparable to Garnham et al.'s (1995) object condition. It might be that the conditions realized here and in Garnham et al.'s (1995) final experiment are borderline cases. However, both Garnham et al.'s and the present results indicate that even beyond sentence boundaries there is some influence of morphosyntactic information. This is what is predicted by the model.

In sum, the results of Experiment 5 give support to the idea that requirements of language processing determine what is retained in short-term memory. However, this support is limited so far. Further experiments investigating different types of information are needed. Ideas for future studies that could explore this point will be sketched in the General Discussion.

With respect to earlier results on memory for surface information about sentences in texts (e.g., Jarvella, 1971; Sachs, 1967, 1974), the results of Experiment 5 are worth considering in three aspects: the retention of gender information, the occurrence of intrusion errors, and overall recall accuracy. As to gender information, the present results question the strict assumption that only conceptual information is represented as soon as a sentence boundary is reached.

However, Jarvella's assumption that only conceptual information is maintained as soon as a sentence has been interpreted is inspired by the same idea as the account promoted in this study: namely by the idea that this is the critical information needed for further processing, while the system gets rid of irrelevant information. However, conceptual information is not the only type of information that is needed further on. With respect to intrusion errors, the present results demonstrate that also lexical information is maintained for a longer period of time. In both states of processing and for both types of lure words there were clear intrusion effects observable. That is, participants erroneously substituted the target word in the sentence more often when a previous sentence contained a lure word. Although the lure word had occurred far before sentence recall took place, the lexical information was still available to a high enough degree to interfere with lexical (re-)access.

With respect to participants' general recall accuracy, the results of this experiment do not correspond with Jarvella's results in all aspects. While his results could be replicated in so far as recall performance of the critical penultimate clause

was influenced by its state of processing (i.e., it was recalled better when it belonged to the most recent sentence than when it belonged to the previous sentence), this was not the case for the comparison between recall accuracy for the penultimate and for the ultimate clause. In contrast to Jarvella's experiments, there was no general advantage for the most recent clause. Instead, the penultimate clause was recalled more accurately than the ultimate clause, irrespective of whether or not they belonged to the same sentence. A first reason that comes to mind would be the recall cue that was provided for the first clause and that enabled close to perfect recall for this clause. This might also have served as a retrieval cue for the second clause. However, Jarvella (e.g., 1971, Exp. 2) replicated his results with a cueing condition (similar to the one applied here). Moreover, in the present experiment recall for the second clause was better when it belonged to the final sentence and not when it formed a sentence with the cued clause. This should not have been the case if the advantage for the second clause was due to the benefit from the first clause. Yet there is a potential confound in the comparison of the second and the third clause. Different clauses with different wording, different syntactic structures, etc. were compared here. The materials were not designed to achieve the highest possible similarity between the two clauses. Therefore, any difference in recall accuracy might be caused by differences in the complexity of the clauses. Anyway, this is not the first time that a recall advantage for the final processing unit could not be observed. Engelkamp and Rummer (2002) report similar results to the ones obtained here. The authors had their participants repeat two unrelated sentences that were presented auditorily. As in the present experiment, recall for the first of the two sentences was more accurate than for the second sentence. However, it cannot be concluded why the recall advantage for the final clause could not be observed here.

In general, one can state that this task was far more difficult than the single-sentence recall task. This is reflected in the low proportion of correctly reproduced target words (72% overall) and in the high proportion of target omissions (about 12% overall), even if the items that led most frequently to recall failure were excluded. In addition, the results of this experiment were influenced by inter-item variation to a higher degree than the sentence recall experiments. The materials included items for which no intrusion was observed and items with a high proportion of intrusions.

9. General Discussion

After a brief overview of the main findings this General Discussion will be divided into three main parts. I will successively address the implications of the present findings for the two fields of research within which the topic of this thesis lies: namely, (1) short-term memory and (2) psycholinguistics. This thesis ends with (3) a few final words and a sketch of ideas for subsequent research.

Overview of the findings. The main purpose of this study was to provide further evidence for the assumption that short-term retention of verbal materials strongly depends on the language processing system. This issue was addressed by demonstrating an effect in a short-term memory task that also occurs in language perception and language production. Furthermore, the effect cannot be explained by models of working memory that do not or only minimally take into account characteristics of the language processing system. More specifically, it was tested whether grammatical gender does not only influence processing but also retention of linguistic materials. A gender congruency effect manifests itself in slower reaction times after a gender incongruent as compared to a gender congruent prime in language perception and in slower reaction times in the presence of a gender incongruent as compared to a gender congruent distractor word and more gender congruent than gender incongruent errors in language production. Similar influences of gender congruency could in fact be observed with respect to semantic substitution errors in *recall* of syntactically structured materials.

The present findings corroborate previous evidence for a close connection between long-term memory and short-term memory in the field of language. These sentence and text recall data demonstrate the relevance of using sufficiently complex materials and allow for stating (1) that (certain) morphosyntactic representations contribute to short-term retention of sentences (Experiments 1 & 2), (2) that such a contribution is not restricted to immediate recall (Experiment 3), (3) that it is independent of whether or not the morphophonological specification of the morphosyntactic feature is unambiguous (Experiment 4), and (4) that the involvement of grammatical gender information is not influenced by clause recency but extends to retention of short texts (Experiment 5). This final result points beyond the mere fact that an influence of gender information exists – which can most straightforwardly be explained in

terms of language-based working memory models – by providing first, preliminary evidence for the assumption that characteristics of language processing determine what type of information will be retained in short-term memory and for how long such a short-term memory representation will be maintained.

The main purpose of the present study was to further incorporate psycholinguistic theories into models of short-term memory. Thus, in the following subchapter, the results will be discussed with respect to their possible implications for (particularly language-based) models of short-term memory.

9.1 Implications for models of short-term memory

I have argued that one needs to incorporate models of language comprehension and language production to account for verbal short-term memory performance adequately. This should in particular be the case if more complex syntactically structured materials as sentences or texts are considered. In the following paragraphs, I will address the question of how the results from this study can be accounted for by the models described in Chapter 1. Before coming to language-based models, I will briefly discuss the results in the context of Baddeley's and Cowan's memory models.

The latest version of Baddeley's working memory model (e.g., 2000, 2003) has incorporated a new component that is meant to account for long-term memory influences on short-term memory tasks: namely, the episodic buffer. One central type of long-term memory information that comes into play as a consequence is linguistic information. With respect to verbal retention, the episodic buffer integrates information stored in the phonological loop with different types of information from the language processing system. What kind of information gets into the episodic buffer and how the representations that have entered it get together to form a common episodic trace has not yet been specified. Therefore, in the present state of Baddeley's (2000) working memory model, nearly any information can be regarded as entering the episodic buffer so that the model can account for almost any data. The results from the present study do not just reflect the additional influence of one more type of information. Instead, the availability of grammatical gender information resulted in reduced priming effects of semantically related nouns, which demonstrates the mutual influence and interdependency of these different information types. Although the present data do not falsify Baddeley's (2000) model, it is clear that the results obtained here

confirm the predictions of language-based models of verbal short-term memory. Furthermore, it is clear that the episodic buffer theory is less precise in this respect and thus less preferable.

A second more general framework for working memory that has been introduced above is the tripartite memory system by Nelson Cowan (e.g., 1995, 1999). Its emphasis is put on the focus of attention. This limited-capacity system is assumed to be able to simultaneously hold four items that are consciously attended to. In most working memory tasks, other mechanisms add to the retention of items in the focus of attention so that more than four items can be retained. Particularly, Cowan assumes that parts of long-term memory are pre-activated and that they are therefore of higher accessibility than the rest (but of lower accessibility than items in the focus of attention). In addition, items in the focus may point to certain entries in long-term memory (e.g., via associations), which can thus be accessed more easily than other items in long-term memory, although they need not be pre-activated. The interaction of these representations that differ in their current degree of activation constitutes performance in working memory tasks.

The present experiments have in particular explored processes in the pre-activated part of long-term memory. To explain sentence recall in Cowan's nomenclature, one can assume that what is stored in the focus of attention are propositions derived during sentence comprehension (see also Kintsch & van Dijk, 1978, for the assumption that about four propositions can be stored in short-term memory). Each proposition would be connected with words (in long-term memory) that can optimally be used to express the respective proposition. This can potentially account for spontaneous intrusions of better fitting words. Those lexical entries and syntactic structures that have been used in the presented sentence should be pre-activated and thus be most easily accessible. Such an explanation is very closely related to Potter and Lombardi's (1990) conceptual regeneration hypothesis.

The assumption that propositions are stored in a limited-capacity system is supported by findings from dual-task experiments. An external memory load can decrease sentence comprehension performance and this negative effect increases with the number of propositions in the sentence, independent of the number of words (for an overview see Caplan & Waters, 1999). This fits with the limited capacity of the focus of attention. If the number of propositions in a sentence exceeds the num-

ber of chunks that can be stored within the focus of attention, the information cannot be integrated as completely as if all propositions were represented in the focus of attention. Cowan's model is not as elaborate with respect to prose recall as it is in other aspects. Thus there is some scope of interpretation for how to account for the gender congruency effect and similar caveats as for the interpretation in terms of the episodic buffer apply to an interpretation within Cowan's framework. However, the general assumption that activated parts of long-term memory are involved in short-term memory tasks includes the assumption that representations that are part of the language processing system (as a part of long-term memory) are also involved in recall of verbal materials. Hence, the discussion with respect to Cowan's model can take place in combination with the discussion of language-based models of short-term memory. These models elaborate the connection between linguistic long-term knowledge and short-term memory for linguistic materials and can thus be considered language-specific elaborations of parts of Cowan's framework.

The findings from the present experiments are compatible with language-based models of short-term memory. As I have stated in the joint discussion of Experiments 1 to 4 (cf. Subchapter 7.4 above), the present study provides evidence for the assumption that not only phonological and lexical-semantic representations are shared between language processing and short-term memory but also morphosyntactic information. The gender congruency effect in short-term recall of syntactically structured materials has been demonstrated across modalities, across different sets of materials, and across task variations. Since grammatical gender is a linguistic feature that is normally not incorporated in models of short-term memory, these data clearly point to the need of models of verbal recall that are based on language processing. Although the models proposed by Dell et al. (1997) and Martin et al. (1999) are restricted to word processing, word retention, and the contribution of lexical-semantic and phonological information so far, they include the assumption that other types of linguistic representations can play a role.

The account that has been promoted here (see also Subchapter 1.4.3 above) already focuses on retention of sentences and, consequentially, on the contribution of additional types of representations. In addition to focusing on syntactically structured materials, this model suggests that it is language processing (i.e., language comprehension and language production) that *drives* short-term memory. Grammatical gen-

der influences lexical selection in production of syntactically structured utterances (e.g., Marx, 1999; Kulke & Blanken, 2001; Vigliocco et al., 2004) and, as the present results indicate, also influences lexical selection in reproduction of syntactically structured materials.

The influence of gender information is motivated by the need to access gender information in the production of succeeding utterances (e.g., to produce a pronoun that is gender congruent to the noun it refers to) and in comprehension of subsequent utterances (e.g., to determine the correct referent for a subsequently encountered pronoun). As in the context of anaphor processing gender information can be needed across sentence boundaries, it was suggested that a gender congruency effect should be observed no matter if there is an additional sentence boundary between the target word and recall of the respective sentence or not. Experiment 5 suggests the existence of such an effect. With some caveats in mind, it can be concluded that a joint representation of the meaning and the grammatical gender of a noun in a text is still available even after an intervening sentence. And it can definitely be concluded that an influence of grammatical gender is not restricted to the task of recalling single sentences.

In the description of the account on which my study is based (see Subchapter 1.4.3 above), I have sketched the processes that a sentence undergoes from comprehension to regeneration. In the following paragraphs, I want to tie in with this and give an account of how the higher proportion of gender congruent as compared to gender incongruent lexical competitors can be accounted for in this model. To do so, I will draw on theories of language processing, focussing on language production.

In the course of sentence comprehension, propositional-conceptual representations are built and lexical entries and syntactic frames are activated (e.g., Friederici, 2002; Gibson & Pearlmutter, 1998). The frequent observation that the syntactic structure of a perceived sentence can prime syntactic structuring of subsequent sentence production makes it plausible to assume that syntactic frames are shared between comprehension and production (e.g., Branigan, Pickering, & Cleland, 2000; Potter & Lombardi, 1998). Some of these data point to a long-term, implicit learning account of syntactic priming (for a recent overview see Chang, Dell, & Bock, 2006). For instance, Bock and Griffin (2000) demonstrated syntactic priming over a delay of up to ten sentences between prime and target and independent of a meaning overlap.

However, Pickering and Branigan (1998; see also Cleland & Pickering, 2003) claimed for a second type of syntactic priming that lasts a shorter period of time and that can be influenced by lexical overlap and semantic similarity between prime and target (for a discussion of multiple underlying cognitive bases of structural priming and an overview of both accounts of syntactic priming see Ferreira & Bock, *in press*). The present findings seem to apply rather to the latter type of priming, since meaning overlap plays an important role in the intrusion paradigm and since there were no intrusions of other words presented on the distractor list, some of which were also gender congruent with the target word and could thus have benefited from shared gender-specified syntactic frames. In addition, if such frames were subject to long-term learning, independent of the meaning they are associated with, in the course of processing a few sentences, all types of gender-specified frames should have been activated and thus interfere with each other during sentence (re-) production.

The data of this study suggest that such gender-specified frames for noun phrases (activated in sentence comprehension) are retained over the short-term, stay associated with the respective meaning, and can thus influence sentence reproduction. When a sentence is regenerated, starting from a conceptual representation, previously activated syntactic frames have a higher probability of being used for production (see also Pickering & Branigan, 1998). What follows has to be accounted for differently whether or not a model allows for activation flow from syntactic frames to lexical representations. In the following paragraphs, the observed pattern of data will consecutively be interpreted in terms of both types of models.

Interactive activation models generally assume bidirectional connections between levels (such as word forms and lexical semantics, e.g., Dell, 1986; Dell et al., 1997, see also Chapter 1.3). Although these accounts are less specific with respect to interaction between syntactic frames and lexical representations, one can assume that syntactic frames feed back to the lexical level and thus influence lexical retrieval. Due to this feedback, activation should increase for those lexical candidates that fit the syntactic constraints and the activation might then be sufficient to erroneously select non-target representations that are already highly activated because of conceptual input and lexical priming (for a possible interaction between syntactic frames

and lexical representations and a lexical-semantic buffer, respectively, see also Vigliocco, Lauer, Damian, & Levelt, 2002, and Martin & Freedman, 2001).

If the gender incongruent lure word has been primed as part of the word list, both lure and target representations in the mental lexicon should be highly activated. During reproduction of the sentence the lure word receives more activation from the propositional representation(s) than the target word. Yet the target receives extra activation from the gender-specified syntactic frame that has been activated in the course of comprehension. This activation has to be maintained till reproduction (either in a separate buffer along the lines of the model proposed by Martin et al., 1999, or as prolonged activation within the language processing system along the lines of the model proposed by Dell et al., 1997). The gender-specified frame activates the target but not the gender incongruent and thus morphosyntactically incompatible lure.

In contrast, if the gender congruent lure has been presented, its representation in the mental lexicon is already highly activated because of lexical priming and the activation from the propositional level. In this case, activation from the previously activated syntactic frame flows to the lexical representations of both the target and the gender congruent lure because both are compatible. This makes erroneous selection of the gender congruent lure more probable than selection of the gender incongruent lure. If the gender congruent lure word has not been primed, the additional syntactic activation should not matter that much as its level of activation in the mental lexicon is substantially lower. This is reflected in the comparable (low) spontaneous intrusion rates for gender congruent and gender incongruent lure words. Vigliocco et al.'s (2004) language production data suggest that activation from syntactic frames alone does not result in a gender congruency effect but that additional feedback from morphophonologically specified forms is necessary. As in the final experiment the morphophonological form of the determiner did not influence the gender congruency effect, this type of information seems to play only a minor role in short-term retention.

Other language production models do not allow syntactic frames to influence lexical selection. The most prominent of these models is WEAVER++ (Levelt et al., 1999). As in this model neither activated gender-specified frames nor gender nodes can activate lexical representations, the gender congruency effect in short-term sen-

tence recall cannot be explained within the production system itself. Instead, the effect has to be attributed to the monitoring processes, which supervise language production and are able to detect errors and initiate repair of these errors even before they have been produced (i.e., covert repairs). An error is the more likely to be detected the more target and error differ. If the target gender and/or the gender-specified frame are still active during production, the gender discrepancy could result in more covertly repaired gender incongruent intrusions and more undetected gender congruent intrusions. According to this explanation, spontaneous incongruent intrusions should be more likely to be detected by the monitor as well and should therefore be less frequent than spontaneous gender congruent intrusions. This was not the case in the experiments of this study, but due to the overall small number of spontaneous intrusions the present data are not very decisive in this respect.

As a sidestep at the end of this subchapter, one can also consider further implications of the view presented here – apart from strict mechanisms of language processing. In addition to automatic requirements of language processing, one can assume that intentions of the listener or speaker influence what kind of information is retained. In line with this are findings that surface information is more accurately remembered for personal or highly interactive statements (Keenan, MacWhinney, & Mayhew, 1977) or for jokes and sentences with sarcastic meaning (Murphy & Shapiro, 1994). In these cases it is not sufficient or at least not satisfactory to keep in mind just a rough gist representation. Instead, jokes often need to be remembered literally for preserving the punch line and the question whether a statement is ironic rather than serious often depends solely on the slightest difference in intonation. In these contexts it is thus functional to focus on surface information in addition to conceptual information. While the present work has concentrated on rather automatic linguistic processes, the general account on which this study is based provides the opportunity to account for a wide range of phenomena.

In the next subchapter, I will turn to the second main pillar of my study: namely, models of language production and language comprehension. Since the working memory account that I have presented largely depends on such models and since working memory also plays a role in some models of language comprehension, I will discuss implications of the present findings for such models in the following.

9.2 Implications for models of language production and language comprehension

The main focus of my study was on the elaboration of models of short-term memory. Nevertheless an additional theoretical issue can be addressed. The presupposition that verbal short-term memory is based on language processing implies the assumption that results from a short-term memory task can enlighten the processes of language production and language comprehension in some way. This is particularly the case for the following two research areas: (1) what has been investigated here were induced errors of lexical access so that it suggests itself to discuss the present findings with respect to theories of lexical selection (and in particular with respect to the role of grammatical gender in lexical selection); (2) what is more, several theories of sentence processing include working memory as a construct.

Therefore, two aspects will subsequently be discussed in this subchapter: the first one concerns possible implications for the explanation of gender congruency effects in speech production and the second one concerns potential implications for the incorporation of the concept “working memory” in models of sentence comprehension.

Grammatical gender and lexical selection. Although the present study did not aim at testing models of language comprehension or language production, the proposed connection between language processing and short-term memory suggests discussing the gender congruency effect in short-term sentence recall with reference to such models. Gender congruency effects in genuine psycholinguistic studies predominantly occurred in (minimal) syntactic contexts. This was also the case for the observation of a gender congruency effect with (induced semantically related) errors in sentence recall. Intrusion errors are assumed to arise in the regeneration stage of recall rather than in comprehension (see also Lee & Williams, 1997; Potter & Lombardi, 1990, 1998) and the dependent variable of the present experiments was a production rather than a comprehension variable so that little can be said about implications for gender effects in comprehension so far. Thus, I will focus on the theories of gender influences on lexical selection in language production.

Gender congruency effects in speech production have been explained in the context of different models and with respect to different processes. The main difference between these accounts is that gender congruency effects are either viewed as

resulting from interference at the selection of a *noun's gender* (e.g., Schriefers & Teruel, 2000) or from interference at the selection of a *determiner form* (Schiller & Caramazza, 2003). A critical finding for these contrasting theories is the absence of a gender congruency effect in bare noun production. In the first account, this is explained by assuming that gender is activated only in syntactically framed contexts, otherwise it is blocked or bypassed. This is also supposed to account for the absence of a gender congruency effect in the production of German and Dutch plural noun phrases (Schiller & Caramazza, 2003). In this context, one has to assume that gender does not have to be selected in the plural because in these languages gender is never necessary for agreement with plural forms, neither for determiner nor for pronoun production (see also footnote 1; but see Janssen & Caramazza, 2003; Schriefers, Jescheniak, & Hantsch, 2002, 2005, for data suggesting that gender is accessed in the production of German plural noun phrases and that determiner forms compete for selection).

The latter account assumes that gender nodes are only activated and not selected. Therefore competition for selection cannot occur at this stage but is shifted to selection of determiner forms. Assuming cascaded processing, the activated gender node then sends activation to its determiner node which competes for selection with other determiner nodes. In the presence of a distractor, the lexical representations of both target and distractor noun spread activation to their gender nodes, which in turn activate their determiner forms. Faster lexical selection in the presence of gender congruent distractors as compared to gender incongruent distractors should therefore be due to determiner selection. As there are no competing plural determiner forms in Dutch and in German, gender congruent and gender incongruent distractors should not lead to different naming latencies for plural noun phrases, nor should this be the case if no determiner has to be produced at all.

With respect to this controversy, particularly the results of Experiment 4 are relevant. Interestingly, these findings differ from results of language production studies in their independence of morphophonological determiner specification. A gender influence on semantic errors in Vigliocco et al.'s (2004) study could be observed when there was a determiner mismatch between intended and produced items but was not found when gender incongruent errors had the same determiner form. Vigliocco et al. (2004) put forward the thesis that an influence of gender-specified syn-

tactic frames was not strong enough to produce a gender congruency effect. In other words, an additional influence of a morphophonologically specified determiner was required. This was not the case in my sentence recall experiment. It seems as though, in contrast to language production, gender contributes to verbatim sentence recall without support from the determiner form.

This discrepancy between the production of noun phrases in picture naming and the present experiment can be attributed to the fact that Experiment 4 was a short-term memory experiment (as were the other experiments in this study). Hence, speech production was only one part of the recall task applied here. In picture naming, the speech production system gets its input solely from the conceptual system, which is activated by the pictorial stimulus. In contrast, the input in a sentence recall task is the presented sentence, that is, input from the speech comprehension system. The sentence production process is influenced by components that have been activated during speech perception and are retained until (re-)production. Taking this into account, the results of the present study cannot be consulted to decide whether gender or determiner selection is the basis of the gender congruency effect in lexical access tasks. What can be said is that the results do not support the assumption that it is only the morphophonologically specified determiner that counts when it comes to gender congruency. However, although the present data are more in line with gender selection accounts, they do not rule out a determiner selection explanation for speech production, since probably perception, retention, and production contribute to the present effect, which thus cannot be mapped onto lexical selection in production only.

A second area in psycholinguistics for which the results of this study can be consulted is the subject of working memory involvement in sentence comprehension. This issue will be addressed in the following section.

Working memory in language comprehension. Many theories (some of which will briefly be described in the following) assume that working memory is involved in language comprehension as a constraining factor. Obviously, the incorporation of concepts of short-term memory into psycholinguistic theories is a relevant application for language-based models of short-term memory. Where most theories of short-term memory include *language* only in a very abstract way (or restrict language

mainly to phonology), the same holds true for various theories of sentence processing with respect to *working memory* or *working memory load*. Although in several of these theories working memory is included as something that constrains processing, it is not specified what kind of working memory is involved in sentence comprehension. As Gordon, Hendrick, and Levine (2002, p. 426) state, “researchers interested in sentence processing have been agnostic as to how their proposed processing mechanisms map onto possible specializations of working memory”. Such theories can benefit from working memory theories that take into account several representations relevant for language processing.

A working memory based theory of sentence comprehension is Gibson’s Dependency Locality Theory (1998). In this account, two important factors that determine the complexity of a sentence are related to working memory: first, costs associated with keeping track of expected syntactic elements are postulated (*storage costs*, e.g., Chen, Gibson, & Wolf, 2005); second, costs are proposed that are associated with connecting an incoming word to the structure that has been built thus far, including matching earlier expectations (*integration costs*). Integration cost has been demonstrated to be sensitive to distance, in such a way that longer distance integrations give rise to greater complexity in terms of prolonged reading times (Gibson, 1998, 2000; Grodner & Gibson, 2005). For both processing operations (prediction storage and integration), elements need to be maintained and this is assumed to take place in working memory.

What exactly the involved working memory component is is not specified. On the basis of results from dual-task experiments, other theorists have argued that working memory is not involved in syntactic processing (e.g., Waters & Caplan, 1999). However, the circumstance that memorizing a digit list does not interfere with interpreting syntactically complex sentences only indicates that the mechanism that supports digit retention is not involved in syntactic processing. Recently, Gordon et al. (2002) argued that the load manipulation applied in these dual-task experiments did not affect the verbal working memory resources that contribute to sentence comprehension. Whether a working memory task interferes with a linguistic task should depend rather on the representational nature of the memory load and its relation to the representational nature of the to-be-comprehended linguistic materials. If the items that must be retained in a concurrent task are very similar to the items that

must be kept active in memory during the comprehension process, interference should be greater than with a common digit retention task.

Gordon et al. (2002) introduced a novel dual-task paradigm, where participants read sentences that were either of high or of low syntactic complexity and which contained either occupations (e.g., “It was the dancer that liked the fireman before the argument began” vs. “It was the dancer that the fireman liked before the argument began”) or personal names (e.g., “It was Tony that liked Joey before the argument began” vs. “It was Tony that Joey liked before the argument began”). Concurrently, participants had to remember a list of three words, which could also be either occupations (e.g., poet, cartoonist, voter) or personal names (e.g., Joel, Greg, Andy). Thus, two match conditions (memory-nouns and sentence-nouns from the same category) could be compared with two non-match conditions (memory-nouns and sentence-nouns from different categories). After each sentence, participants answered a comprehension question about the sentence and recalled the word list.

As hypothesized by Gordon et al. (2002), the similarity between the memory-nouns and the sentence-nouns affected comprehension of the more complex sentences to a larger extent. They argued that in cases in which the memory representations of the nouns on the list were similar to the memory representations of the antecedent of the verb, interference took place, in such a way that it was harder to identify the antecedent among the representations available in short-term memory. Since antecedent identification was easier in the less complex sentences, the similarity manipulation affected the syntactically complex sentences to a higher degree. With a different type of syntactic complexity variation, Fedorenko, Gibson, and Rohde (2006) obtained similar results with respect to reading times.

Increasing interference between short-term memory and sentence processing tasks with similar materials has also been observed for phonological similarity. When participants had to read sentences with repeated phonemes, reading times were longer when a list of phonologically similar items had to be maintained meanwhile (with each word on the list starting with the same phoneme that was repeated in the sentence; McCutchen, Bell, France, & Perfetti, 1991). In contrast, phoneme repetition in sentences did not influence participants’ ability to resolve syntactic ambiguities (Kennison, 2004).

As these results indicate, it is necessary for theories of sentence comprehension to incorporate a notion of working memory that distinguishes retention of different types of representations (phonological, semantic, syntactic, etc.), which suggests that the working memory concept is very closely linked to language processing. Assuming that syntactic, semantic, and phonological representations are stored separately, they can thus contribute to different tasks separately. Maintaining a representation of what syntactic elements are still needed to complete the current sentence demands different short-term memory representations than retaining a digit list or than keeping in mind the essence of what has been said so far. This is the basis of language-based models of short-term memory. It is thus recommendable to apply working memory tasks that imply representations similar to those that are relevant for sentence comprehension in order to investigate working memory influences on sentence comprehension. The present results provide additional support for the assumption that similarity between the to-be-remembered and the to-be-processed items increases interference. Error rates were highest for those competitors that were semantically *and* morphosyntactically similar to a “Sollbruchstelle” in the sentence.

The accounts sketched above indicate that theories of sentence comprehension start specifying what kind of working memory contributes to what kind of processing operation instead of using working memory or working memory load as an empty phrase that serves to “explain” processing difficulties that cannot be accounted for within the language processing system. Kennison (2004, p. 506) interprets her findings as “preliminary evidence, supporting a view of working memory in which processing involved in maintaining the phonological representations of words in the sentence is carried out separately from processing involved in other aspects of sentence comprehension, including the resolution of syntactic ambiguity.” And Gordon et al. (2002, p. 430) argue that “the special characteristic of language comprehension that supports efficient retrieval of linguistic information from memory is the ability to quickly generate very fine-grained representations of utterances that are richly differentiated and highly organized. This ability greatly reduces the degree of memory interference that arises during sentence comprehension compared with simply trying to remember a list of unrelated words, and provides the basis for the generally high level of memory performance observed with linguistically coherent material, a

high level of performance that can, however, be strained by certain types of syntactic complexity.“

Although the results from the present study cannot account for syntactic complexity phenomena, the general idea of verbal short-term memory as being closely linked to the different representational levels of the language processing system seems to provide a good basis for a more specific use of working memory as one constraining factor in sentence processing.

In the final subchapter, I will end with some remarks on lessons that have been learned and with an outlook on prospects for future work, including a consideration of some drawbacks of this study.

9.3 *Final words and future work*

Altogether, the present study provides evidence for a close connection between language processing and short-term memory with respect to an influence of morphosyntactic information on verbatim sentence and text recall. In addition to the observation of syntactic priming in terms of what has been perceived is most probable to be produced (Lombardi & Potter, 1992; Potter & Lombardi, 1998), the gender congruency effect suggests that pre-activated information at a morphosyntactic level can reduce the impact of priming at the lexical-semantic level. In my opinion, this morphosyntactic contribution goes beyond adding just one more influencing factor to verbal working memory. The general idea was that information that may be needed for subsequent language processing is retained in short-term memory and that what is retained in short-term memory is information that may be needed further on. Such an integrated account of verbal short-term memory and language processing requires a focus on syntactically and conceptually structured materials, that is, on sentences or even texts. Hence, in line with other studies that promote language-based models of verbal short-term memory, the present findings demonstrate how important it is to consider (all aspects of) *language* when investigating *short-term memory for language*.

However, the experiments reported here still leave a variety of questions unanswered. They provide first insight into some aspects of verbal short-term memory but do not allow to falsify certain models or to decide whether the model that motivated this study can indeed account for the findings best. If one follows Rummer and

Schweppe's account, additional predictions can be made. I will sketch aspects that should be investigated in future studies in the rest of this subchapter.

Future prospects

There are different directions in which the question of the interdependence of language processing and verbal short-term memory can be addressed in future studies.

(1) Further types of information with different processing characteristics need to be investigated, in particular those that are required for initial sentence comprehension but not for subsequent processes.

(2) It has been posited that processing of sentences differs from processing of lists. Based on the assumption that processing characteristics determine retention characteristics, different hypotheses concerning the role of grammatical gender information apply to list retention.

(3) Evidence for the relation between language processing and short-term retention comes also from neuropsychological patients who show comparable deficits in both kinds of tasks. Similar results should be obtained for processing and retention of grammatical gender.

(4) In addition, it needs to be specified how the language dependency of verbal short-term memory emerges.

In the following paragraphs, I will successively address these points in order to determine starting points for future studies.

The role of phonological information and syntactic number information. A possible implication of the account on which this study was based is that one would expect different results from those obtained here for linguistic representations that are *not* required for subsequent discourse processing although they are relevant for initial sentence processing. An influence of phonological information should be restricted to the most recent sentence, since after initial processing steps maintenance of a lexical representation should be sufficient for producing or comprehending further parts of discourse. Short-term memory data on this point are not yet available, but psycholinguistic data demonstrate that *overt phonological* gender specification has additional impact on anaphor resolution only for the most recent clause, indicating that

the contribution of a phonological short-term representation is restricted to the most recently encountered informational unit (Meyer & Bock, 1999).

Similar conclusions can be drawn for other types of morphosyntactic information such as syntactic number information. In contrast to grammatical gender, number is closely related to conceptual information. Yet there are some cases in which grammatical and conceptual number diverge (e.g., in collective nouns such as audience, swarm, or team). Since this is only very rarely the case, it is normally not necessary to maintain an extra representation of grammatical number in addition to a conceptual number representation. This suggests that conceptual number should be represented for a longer period of time but grammatical number should not. This could be tested by comparing pronouns that match a collective noun antecedent either in conceptual number (i.e., plural pronouns) or in grammatical number (i.e., singular pronouns). Reading times should differ for these two agreement conditions depending on the position of the pronoun. Longer reading times indicate processing difficulties. With short distance, pronouns that match the antecedent in grammatical number should be processed faster, whereas with longer distance, pronouns that match the antecedent in conceptual number should be processed faster. This would function as an indirect measure of whether or not grammatical number is prominently represented in short-term memory, since only if this were the case could it override conceptual number agreement.

Grammatical gender in list recall. Another way to further test the assumption that verbal retention is based on the requirements of language processing is to change the to-be-recalled materials. List processing differs from sentence processing so that, in consequence, list recall should similarly differ from sentence recall. For neither processing nor recall of a word list is it essential to extract gender information. While phonological differentiation is necessary to discriminate between different words on a list (particularly in lists consisting of phonologically similar words such as “man, cat, map, cab, can”), gender information can – in rare cases – be needed to discriminate between word meanings in a sentence context (e.g., “Meine Oma liebt die Heide_[fem].” (My grandma loves the heath.) vs. “Meine Oma liebt den Heiden_[masc].” (My grandma loves the heathen.)). However, it would not help list recall. To-be-recalled lists normally do not contain determiners so that a list including both mean-

ings of “Heide” would just be a list with the word “Heide” repeated. As it is predicted that only relevant information will be retained, the insignificance of grammatical gender information in single word or unrelated list processing suggests that it does not influence list recall the way it influences sentence recall, unless a list is made up of noun phrases including determiners. Further research is needed to clarify this point.

However, the assumption that a gender congruency effect should not be observable or should at least be attenuated with one of the most important tasks in research on verbal short-term memory questions the promoted importance of the present results. Nevertheless, beyond the tendency to regard one’s own results as milestones in science, I do not think that the results are irrelevant because of this – theoretically based – limitation. Research on serial recall of lists of digits, letters, non-words, and words can still be conducted and interpreted without agonizing about grammatical gender. Yet the applicability of theories based on list recall to tasks that require processing of more complex materials (such as problem solving or language comprehension which normally imply processing of connected discourse) is indeed questioned. It is questioned the more evidence is found for theories that can account for retention of more complex materials. These theories postulate recall influences of information types that are not (or solely to a smaller degree) provided by digits, letters, or single words and thus make possibly different predictions for the complex task of interest. This is particularly the case if such additional representations do not only add to the influence of other representations but counteract the influence of well-known representations (as has been observed in the interaction of lexical-semantic and grammatical gender information in the present study). With respect to consequences for the application of working memory theories to language comprehension, this has been discussed in the previous chapter. Generally, it seems worth broadening the scope of verbal recall and the inclusion of theories of language processing allows for predicting what kind of information is represented in short-term memory under what circumstances.

Neuropsychological patient studies. Evidence for the involvement of phonological and lexical-semantic representations in short-term retention comes not only from experiments with unimpaired participants, but in addition from neuropsychological

patient studies (e.g., Martin & He, 2004; Martin et al., 1999, see also Chapter 1.3 above). This should also be a way to further test the assumption that grammatical gender information is shared between sentence processing and sentence retention. A first indicator is the performance of the French patient CB (Alario & Cohen, 2004, see also Subchapter 1.4.3 above). He had problems producing function words in written sentence production and reproduction. In particular, the grammatical gender of determiners was problematic for him. A modified version of the intrusion paradigm has already been applied to patient studies (Rummer et al., *subm.*). The gender variant of this paradigm could also be used in this context. Patients like CB or other aphasic patients who have difficulties producing determiners should then show an attenuated or even eliminated gender congruency effect.

How can requirements of language processing influence short-term retention characteristics? Despite its appeal, Rummer and Schweppe's model so far lacks an elaborate account of how the requirements of language processing 'decide' what kind of information is needed and how they cause the information that is needed to be maintained. To avoid a homunculus interpretation of this idea, it is necessary to further specify the processes of verbal retention, language processing, and their relation.

One general idea of how requirements of language processing could form other cognitive processes and how therefore differences between tasks and types of representations could emerge is associated with learning-based connectionist models (e.g., MacDonald & Christiansen, 2002). In learning-based connectionist models, frequent use can strengthen the connections between elements. "In spreading activation models of production, practice is assumed to exert its effects by strengthening connections" (Schwartz, Saffran, Bloch, & Dell, 1994, p.72). The stronger a connection, the more activation spreads from one element to another and the higher these elements are activated. Thus, strong connections increase accessibility of the connected entities as compared to weaker connections and allow for easier comprehension or production. The more often one element is jointly accessed with a related element, the stronger the connection between these two elements becomes. For instance, the dates of one's day of birth and month of birth are often accessed together so that producing your day of birth activates the representation of your month of birth to a certain degree, even if this is not required. Similarly, it can be assumed that

upon reading texts and encountering pronouns, one learns that access to a combined representation of both the meaning and the grammatical gender of the previously perceived noun phrases can be helpful for interpreting the pronoun. Thus the connection between the meaning and the gender of a noun needs to be strong enough to allow for later access. The more often one makes use of the connection between a pronoun and the gender information of earlier discourse entities, the stronger such links become and the longer activation can be maintained.

At least two predictions follow from the assumption that experience with language (processing) forms certain characteristics of verbal short-term memory: since strengthening of connections depends on learning, there should be (1) developmental differences in the use of certain types of representations depending on the level of language development (e.g., in phases where grammatical development is not yet advanced children should retain phonological information better than morphosyntactic information, in contrast to adults) and (2) language-specific differences according to the requirements of the *respective* language (e.g., between languages that rely predominantly on conceptual gender such as English, languages with an abstract gender system in which gender markers on nouns are less overt such as German, and languages in which gender markers are more overt such as Italian). Language-specific differences might be observed with respect to anaphoric relations as in example (15) above. Interpretation of that anaphoric pronoun relies to a good degree on gender agreement between pronoun and antecedent. In English (as can also be seen in the translation of the example), this information is not provided. Thus, such anaphor types should be less frequent in English than in German because they are a lot more difficult to resolve. Systematic analyses of speech corpora could provide evidence for this assumption.

There is another problem that is related to the underspecification of the mechanisms that make the requirements of language processing determine what is retained in short-term memory. The argumentation that the language processing characteristics determine the retention characteristics is somewhat recursive, at least when it is combined with theories of language comprehension. To illustrate this problem, let us have a look at psycholinguistic explanations of some problems with anaphoric devices. Pronouns are understood less reliably than nouns when they refer to an antecedent that was mentioned several sentences before. Sanford and Garrod

(1981), for instance, argue that with greater distance between antecedent and anaphoric expression the antecedent is no longer represented in working memory. That is, the difficulty of certain syntactic structures is explained in terms of working memory limitations and it is assumed that the accessibility of representations determines language processing, at least to some degree.

In other words, psycholinguistic theories argue that retention characteristics determine language processing characteristics and we argue that language processing characteristics determine retention characteristics. This can lead to circular explanations of phenomena in language processing and of phenomena in short-term memory. In the worst case, the respective explanations can be broken down into definitions such as “Recursion: see ‘Recursion’”. The problem of circularity becomes particularly evident when psycholinguistic models and models of short-term memory are combined, as has been done here.

The circularity is inherent in the assumption that language and short-term memory are intertwined and thus in the problem itself. It becomes most obvious when one tries to specify what is meant by working memory *and* what is meant by language. As long as one focuses on only one aspect (i.e., either on language processing *or* on working memory) and assumes that the other aspect is a constant, one can circumvent the problem. This makes the circularity less obvious but does not solve the problem either. However, I do not know how to solve this problem that is evidently immanent to the model that I have promoted. This is definitely a shortcoming of the present work and the issue remains to be addressed in the future.

At the end of the day... the study has demonstrated that grammatical gender can significantly influence performance in sentence recall. With respect to text recall, the results point in the same direction but are still less reliable. Whether the account proposed here proves successful and useful with respect to further types of representations remains to be answered in future studies.

References

- Akhutina, T., Kurgansky, A., Polinsky, M., & Bates, E. (1999). Processing of grammatical gender in a three-gender system: Experimental evidence from Russian. *Journal of Psycholinguistic Research*, 28, 695-713.
- Alario, F.-X. & Cohen, L. (2004). Closed-class words in sentence production: Evidence from a modality-specific dissociation. *Cognitive Neuropsychology*, 21, 787-819.
- Allen, R. & Hulme, C. (2006). Speech and language processing mechanisms in verbal serial recall. *Journal of Memory and Language*, 55, 64-88.
- Alloway, T. P. (subm.). What can phonological and semantic information tell us about the mechanisms of immediate sentence recall? Manuscript submitted for publication.
- Altmann, G. T. M. (1998). Ambiguity, parsing strategies, and computational models. *Language and Cognitive Processes*, 3, 73-98.
- Altmann, G. T. M. & Steedman, M.J. (1988). Interaction with context during human sentence processing. *Cognition*, 30, 191-238.
- Atkinson, R. C. & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89-195). Academic Press.
- Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity. *Quarterly Journal of Experimental Psychology*, 18, 362-365.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford Univ. Press.
- Baddeley, A. D. (1993). Working memory or working attention? In A. D. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control* (pp. 152-170). Oxford, England: Oxford University Press.
- Baddeley, A. D. (2000). The episodic buffer: a new component of working memory. *Trends in Cognitive Sciences*, 4, 417-423.
- Baddeley, A. D. (2003). Working memory and language: an overview. *Journal of Communication Disorders*, 36, 189-208.
- Baddeley, A. D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105, 158-173.

- Baddeley, A. D. & Hitch, G. J. (1974). Working memory. In G.A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.
- Baddeley, A. D. & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 28–61). New York: Cambridge University Press.
- Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning & Verbal Behavior*, *14*, 575-589.
- Balota, D. A., Cowan, N., & Engle, R. W. (1990). Suffix interference in the recall of linguistically coherent speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 446-456.
- Barnard, P.J. (1985). Interacting cognitive subsystems: A psycholinguistic approach to short term memory. In A. Ellis (Ed.), *Progress in the psychology of language* (Vol. 2, pp.197-258). London: Erlbaum.
- Bates, E., Devescovi, A., Hernandez, A., & Pizzamiglio, L. (1996). Gender priming in Italian. *Perception & Psychophysics*, *58*, 992-1004.
- Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*, 355-387.
- Bock, K. (1989). Closed-class immanence in sentence production. *Cognition*, *31*, 163-186.
- Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, *129*, 177-192.
- Bosch, P. (1988). Representing and accessing focussed referents. *Language and Cognitive Processes*, *3*, 207–231.
- Bourassa, D. & Besner, D. (1994). Beyond the articulatory loop: A semantic contribution to serial order recall of subspan lists. *Psychonomic Bulletin & Review*, *1*, 122-125.
- Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, *75*, B13-B25.
- Caplan, D. (1972). Clause boundaries and recognition latencies for words in sentences. *Perception & Psychophysics*, *12*, 73-76.

- Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22*, 77-126.
- Caramazza, A., & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: Evidence from the 'tip-of-the-tongue' phenomenon. *Cognition*, *64*, 309-343.
- Caramazza, A., Miozzo, M., Costa, A., Schiller, N., & Alario, F.X. (2001). A crosslinguistic investigation of determiner production. In E. Dupoux (Ed.) *Language, brain, and cognitive development: Essays in honor of Jacques Mehler* (pp. 209-226). Cambridge, MA, US: The MIT Press.
- Chang, F., Dell, G. S., & Bock, K. (2006). Becoming Syntactic. *Psychological Review*, *113*, 234-272.
- Chen, E., Gibson, E., & Wolf, F. (2005). Online syntactic storage costs in sentence comprehension. *Journal of Memory and Language*, *52*, 144-169.
- Clark, H. H. & Sengul, C. J. (1979). In search of referents for nouns and pronouns. *Memory & Cognition*, *7*, 35-41.
- Cleland, A. A., & Pickering, M.J. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language*, *49*, 214-230.
- Coltheart, V. (1999a). Phonological codes in reading comprehension, short-term memory, and memory for rapid visual sequences. In V. Coltheart (Ed.), *Fleeting memories: Cognition of brief visual stimuli*. (pp. 181-224). Cambridge, MA: MIT Press.
- Coltheart, V. (1999b). Comparing short-term memory and memory for rapidly presented visual stimuli. *International Journal of Psychology*, *34*, 293-300.
- Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology*, *55*, 75-84.
- Conrad, R. & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology*, *55*, 429-432.
- Conrad, R., & Hull, A. J. (1968). Input modality and the serial position curve in short-term memory. *Psychonomic Science*, *10*, 135-136.
- Costa, A., Sebastián-Gallés, N., Miozzo, M., & Caramazza, A. (1999). The gender congruity effect: evidence from Spanish and Catalan. *Language and Cognitive Processes*, *14*, 381-391.

- Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin*, *96*, 341–70.
- Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford Psychology Series, No. 26. Oxford University Press.
- Cowan, N. (1999). An embedded-processes model of working memory. In: A. Miyake & P. Shah (Eds.). *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge University Press.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87-185.
- Cowan, N., Saults, J. S., & Brown, G. D. A. (2004). On the auditory modality superiority effect in serial recall: Separating input and output factors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 639-644.
- Crowder, R. G. (1967). Prefix effect in immediate memory. *Canadian Journal of Psychology*, *21*, 450–461.
- Crowder, R. G. (1978). Memory for phonologically uniform lists. *Journal of Verbal Learning and Verbal Behavior*, *17*, 73-89.
- Cubelli, R., Lotto, L., Paolieri, D., Girelli, M., & Job, R. (2005). Grammatical gender is selected in bare noun production: Evidence from the picture–word interference paradigm. *Journal of Memory and Language*, *53*, 42–59.
- Daneman, M. E. & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450-466.
- Daneman, M., E. & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 561–584.
- Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, *93*, 283-321.
- Dell, G. S. & O’Seaghda, P. G. (1992). Inhibition in interactive models of linguistic selection and sequencing. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory and language* (pp. 409-453). San Diego, CA: Academic Press.
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1996). A connectionist model of naming errors in aphasia. In J. Reggia, R. Berndt, & E.

- Ruppin (Eds.), *Neural modeling of cognitive and brain disorders*. (pp. 135-156). New York: World Scientific.
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in normal and aphasic speakers. *Psychological Review*, *104*, 801-838.
- Ellis, A. W., & Young, A. W. (1988). *Human cognitive neuropsychology*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Emerson, M. J., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language*, *48*, 148-168.
- Engelkamp, J. & Rummel, R. (2002). Subordinating conjunctions as devices for unifying sentences in memory. *European Journal of Cognitive Psychology*, *14*, 353-369.
- Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102-134). New York: Cambridge University Press.
- Ericsson, K. A. & Delaney, P. F. (1999). Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 257-297). New York: Cambridge University Press.
- Ericsson, K. A. & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*, 211-245.
- Eysenck, M. W. & Keane, M. T. (2005). *Cognitive psychology. A student's handbook*. Hove: Psychology Press.
- Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. *Journal of Memory and Language*, *54*, 541-553.

- Ferreira, V. S. & Bock, K. (in press). The functions of structural priming. *Language and Cognitive Processes*.
- Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, 25, 348–368.
- Fodor, J. A., (1983). *The modularity of the mind*. Cambridge, MA: MIT/Bradford Books.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14, 178-210.
- Frazier, L., & Clifton, C. (1996). *Construal*. Cambridge, MA: MIT Press.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6, 78-84.
- Friederici, A. D. & Jacobsen, T. (1999). Processing grammatical gender during language comprehension. *Journal of Psycholinguistic Research*, 28, 467-484.
- Garnham, A, Oakhill, J., Ehrlich, M. F., & Carreiras (1995). Representations and processes in the interpretation of pronouns: New evidence from Spanish and French. *Journal of Memory and Language*, 34, 41–62.
- Garrod, S. & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, 8, 8-11.
- Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68, 1–76.
- Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In Y. Miyashita, A. Marantz, & W. O’Neil (Eds.), *Image, language, brain* (pp. 95-126). Cambridge, MA: MIT Press.
- Gibson, E., & Pearlmutter, N. (1998). Constraints on sentence comprehension. *Trends in Cognitive Sciences*, 2, 262–268.
- Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science*, 13, 425-430.

- Grodner, D., & Gibson, E. (2005). Consequences of the serial nature of linguistic input for sentential complexity. *Cognitive Science*, 29, 261–290.
- Grosjean, F., Dommergues, J., Cornu, E., Guillelmon, D., & Besson, M. (1994). The gender marking effect in spoken word recognition. *Perception & Psychophysics*, 56, 590-598.
- Harley, T. (2001). *The psychology of language: From data to theory*. Hove, England: Psychology Press.
- Hemforth, B. & Konieczny, L. (2002). Sätze und Texte verstehen und produzieren. In: J. Müsseler & W. Prinz (Eds.), *Allgemeine Psychologie* (pp. 589-642). Heidelberg: Spektrum Akademischer Verlag.
- Herrmann, T. & Grabowski, J. (1994). *Sprechen: Psychologie der Sprachproduktion*. Heidelberg: Spektrum Akademischer Verlag.
- Hofmann, J. (2006). *Phonological, morphological, and semantic aspects of grammatical gender processing in German*. Dresden: Sächsisches Druck- und Verlagshaus Direct World.
- Howell, D.C. (2002). *Statistical Methods for Psychology*. Belmont, CA: Duxbury.
- Jacobsen, T. (1999). Effects of grammatical gender on picture and word naming: Evidence from German. *Journal of Psycholinguistic Research*, 28, 499-514.
- James, W. (1890). *The principles of psychology*. Henry Holt.
- Janssen, N. & Caramazza, A. (2003). The selection of closed-class words in noun phrase production: The case of Dutch determiners. *Journal of Memory and Language*, 48, 635–652.
- Jarvella, R. J. (1971). Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior*, 10, 409–416.
- Jarvella, R. J. (1972). Speech processing memory. In P.C. Smith (Ed.) *1972 conference on speech communication and processing*. New York: Institute of Electrical and Electronics Engineers
- Jarvella, R. J. (1979). Immediate memory and discourse processing. In: Bower, G. H. (Ed.), *The Psychology of Learning and Motivation* (Vol. 13, pp. 379–421). New York: Academic Press.

- Jefferies, E., Lambon Ralph, M. A., & Baddeley, A. D. (2004). Automatic and controlled processing in sentence recall: The role of long-term and working memory. *Journal of Memory and Language, 51*, 623–643.
- Jones, D. M., Hughes, R. W., & Macken, W. J. (2006). Perceptual organization masquerading as phonological storage: Further support for a perceptual-gestural view of short-term memory. *Journal of Memory and Language, 54*, 265–281.
- Jones, D. M., Macken, W. J., & Nicholls, A. P. (2004). The phonological store of working memory: Is it phonological and is it a store? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 656–674.
- Just, M. A. & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122–149.
- Keenan, I. M., MacWhinney, B., & Mayhew, D. (1977). Pragmatics in memory : A study of natural conversation. *Journal of Verbal Learning and Verbal Behavior, 16*, 549-560.
- Kennison, S. M. (2004). The effect of phonemic repetition on syntactic ambiguity resolution: implications for models of working memory. *Journal of Psycholinguistic Research, 33*, 493-516.
- Kintsch, W. & van Dijk, T. A. (1978). Towards a model of text comprehension and production. *Psychological Review, 85*, 363-394.
- Köpcke, K. M. & Zubin, D. (1984). Sechs Prinzipien für die Genuszuweisung im Deutschen. Ein Beitrag zur natürlichen Klassifikation. *Linguistische Berichte, 93*, 26-50.
- Kulke, F. & Blanken, G. (2001). Phonological and syntactic influences on semantic misnamings in aphasia. *Aphasiology, 15*, 3-15.
- La Heij, W., Mak, P., Sander, J., & Willeboordse, E. (1998). The gender-congruency effect in picture-word tasks. *Psychological Research, 61*, 209-219.
- Lee, M. & Williams, J. N. (1997). Why is short-term sentence recall verbatim? An evaluation of the role of lexical priming. *Memory & Cognition, 25*, 156-172.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1-7.
- Lombardi, L. & Potter, M. C. (1992). The regeneration of syntax in short term memory. *Journal of Memory and Language*, *31*, 713-733.
- Martin, N. & Gupta, P. (2004). Exploring the relationship between word processing and verbal short-term memory: Evidence from associations and dissociations. *Cognitive Neuropsychology*, *21*, 213-228.
- Martin, N. & Saffran, E. M. (1997). Language and auditory-verbal short-term memory impairments: Evidence for common underlying processes. *Cognitive Neuropsychology*, *14*, 641-682.
- Martin, R. C. (1993). Short-term memory and sentence processing: evidence from neuropsychology. *Memory & Cognition*, *21*, 176-183.
- Martin, R. C. & Freedman, M. L. (2001). Short-term retention of lexical-semantic representations: Implications for speech production. *Memory*, *9*, 261-280.
- Martin, R.C. & He, T. (2004). Semantic short-term memory and its role in sentence processing: a replication. *Brain and Language*, *89*, 76-82.
- Martin, R. C., Lesch, M. F., & Bartha, M. (1999). Independence of input and output phonology in word processing and short-term memory. *Journal of Memory and Language*, *41*, 2-39.
- Martin, R. C., Shelton, J. R., & Yaffee, L. S. (1994). Language processing and working memory: Neuropsychological evidence for separate phonological and semantic capacities. *Journal of Memory and Language*, *33*, 83-111.
- Marx, E. (1999). Gender processing in speech production: Evidence from German speech errors. *Journal of Psycholinguistic Research*, *28*, 601-621.
- McCutchen, D., Bell, L. C., France, I. M., & Perfetti, C. A. (1991). Phoneme-specific interference in reading: The tongue-twister effect revisited. *Reading Research Quarterly*, *26*, 87-103.
- MacDonald, M. C. & Christiansen, M. H. (2002). Reassessing working memory: A comment on Just & Carpenter (1992) and Waters & Caplan (1996). *Psychological Review*, *109*, 35-54.

- MacDonald, M., Pearlmutter, N., Seidenberg, M. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676–703.
- Meyer, A. S. & Bock, K. (1999). Representations and Processes in the Production of Pronouns: Some Perspectives from Dutch. *Journal of Memory and Language*, *41*, 281-301.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex frontal lobe tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.
- Miyake, A. & Shah, P. (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. New York: Cambridge University Press.
- Miozzo, M. & Caramazza, A. (1999). The selection of determiners in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 907-922.
- Murphy, G. L., & Shapiro, A. M. (1994). Forgetting of verbatim information in discourse. *Memory & Cognition*, *22*, 85–94.
- Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology*, *78*, 679–684.
- Nespoulous, J. L., Dordain, M., Perron, C., Ska, B., Bub, D., Caplan, D., Mehler, J., & Lecours, A. R. (1988). Agrammatism in sentence production without comprehension deficits: Reduced availability of syntactic structures and/or of grammatical morphemes? A case study. *Brain and Language*, *33*, 273–295.
- Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 411-421.
- Park, N. & Martin, R. C. (2002). A contribution of phonological representations to immediate sentence recall. *15th Annual CUNY Conference on Human Sentence Processing*, New York City, March.

- Patterson, K.E. (1988). Acquired disorders of spelling. In G. Denes, C. Semenza, & P. Bisiacchi (Eds.), *Perspectives on cognitive neuropsychology* (pp. 213–228). Hove: Lawrence Erlbaum Associates Ltd.
- Penney, C. G. (1979). Interactions of suffix effects with suffix delay and recall modality in serial recall. *Journal of Experimental Psychology*, *5*, 507-521.
- Penney, C. G. (1989). Modality effects in the structure of short-term verbal memory. *Memory & Cognition*, *17*, 398-422.
- Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*, 633-651.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*, 1–22.
- Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology*, *10*, 377–500.
- Potter, M. C. (1999). Understanding sentences and scenes: The role of conceptual short-term memory. In V. Coltheart (Ed.), *Fleeting memories: Cognition of brief visual stimuli* (pp. 13-46). Cambridge, MA: MIT Press.
- Potter, M. C. & Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Language*, *29*, 633-654.
- Potter, M. C. & Lombardi, L. (1998). Syntactic priming in immediate recall of sentences. *Journal of Memory and Language*, *38*, 265-282.
- Rummer, R. (2003). Das kurzfristige Behalten von Sätzen. *Psychologische Rundschau*, *54*, 93-102.
- Rummer, R. & Engelkamp, J. (2001). Phonological information contributes to short-term recall of auditorily presented sentences. *Journal of Memory and Language*, *45*, 451-467.
- Rummer, R. & Engelkamp, J. (2003). Phonological information in immediate and delayed sentence recall. *The Quarterly Journal of Experimental Psychology*, *56A*, 83-95.

- Rummer, R. & Engelkamp, J. (in press). Towards a language-based account of verbal working memory. C. Zelinsky (Ed.), *Memory and language*. Amsterdam: Benjamins.
- Rummer, R., Engelkamp, J., & Konieczny, L. (2003). The subordination effect: Evidence from self-paced reading and recall. *European Journal of Cognitive Psychology, 15*, 539-566.
- Rummer, R., Martin, R. C., Schweppe, J., & Bormann, T. (subm.). Phonological information contributes to sentence recall: Support from experiments and neuropsychological patient data. Manuscript submitted for publication.
- Rummer, R. & Schweppe, J. (2005). Evidence for a modality effect in sentence retention. *Psychonomic Bulletin & Review, 12*, 1094-1099.
- Rummer, R. & Schweppe, J. (in prep.). Verbal retention as language processing: A language-based model of short-term memory. Manuscript in preparation.
- Rummer, R., Schweppe, J., & Martin, R. C. (subm.). Oral and written recall of heard and read sentences: Further evidence for the modality congruency hypothesis. Manuscript submitted for publication.
- Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception & Psychophysics, 2*, 437-442.
- Sachs, J. S. (1974). Memory in reading and listening to discourse. *Memory & Cognition, 2*, 95-100.
- Saffran, E. M., Bogyo, L. C., Schwartz, M. F., & Marin, O. S. M. (1980). Does deep dyslexia reflect right-hemisphere reading? In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 381-406). London: Routledge & Kegan Paul.
- Salamé, P., & Baddeley, A.D. (1982). Disruption of memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior, 21*, 150-164.
- Sanford, A. J., & Garrod, S. C. (1981). *Understanding written language*. Wiley: Chichester.
- Schade, U. (1999). *Konnektionistische Sprachproduktion*. Wiesbaden: Deutscher UniversitätsVerlag.

- Scheepers, C. (2003). Syntactic priming of relative clause attachments: persistence of structural configuration in sentence production. *Cognition*, 89, 179-205.
- Schiller, N. O. & Caramazza, A. (2003). Grammatical feature selection in noun phrase production: Evidence from German and Dutch. *Journal of Memory and Language*, 48, 169-194.
- Schmidt, R. (1986). Was weiß der Artikel vom Hauptwort? Ein Beitrag zur Verarbeitung syntaktischer Beziehungen beim Lesen. *Zeitschrift für experimentelle and angewandte Psychologie*, 33, 150-163.
- Schneider, W. & Detweiler, M. (1987). A connectionist/control architecture for working memory. In G. H. Bower (Ed.), *The psychology of learning motivation* (Vol. 21, pp. 53-119). Academic Press.
- Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 841-850.
- Schriefers, H., Friederici, A. D., & Rose, U. (1998). Context effects in visual word recognition: Lexical relatedness and syntactic context. *Memory & Cognition*, 26, 1292-1303.
- Schriefers, H., Jescheniak, J. D., & Hantsch, A. (2002). Determiner selection in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 941-950.
- Schriefers, H., Jescheniak, J. D., & Hantsch, A. (2005). Selection of gender-marked morphemes in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 159-168.
- Schriefers, H., Meyer, A. S., & Levelt, W. J. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, 29, 86-102
- Schriefers, H. & Teruel, E. (2000). Grammatical gender in noun phrase production: The gender interference effect in German. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1368-1377.
- Schwartz, M. F., Saffran, E. M., Bloch, D. E., & Dell, G. S. (1994). Disordered speech production in aphasic and normal speakers. *Brain and Language*, 47, 52-88.
- Schweppe, J. & Rummer, R. (2007). Shared representations in language processing and verbal short-term memory: The case of grammatical gender. *Journal of Memory and Language*.

- Shah, P., & Miyake, A. (1999). Models of working memory: An introduction. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 1–27). New York: Cambridge University Press.
- Trueswell, J. C., Tanenhaus, M. K., Garnsey, S. M. (1994). Semantic influences on parsing: use of thematic role information in syntactic disambiguation. *Journal of Memory and Language*, *33*, 285–318.
- Van Berkum, J. J. A. (1997). Syntactic processes in speech production: The retrieval of grammatical gender. *Cognition*, *64*, 115–152.
- Vigliocco, G., Lauer, M., Damian, M. F., & Levelt, W. J. M. (2002). Semantic and syntactic forces in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 46–58.
- Vigliocco, G., Vinson, D. P., Indefrey, P., Levelt, W. J. M., & Hellwig, F. (2004). Role of grammatical gender and semantics in German word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 483–497.
- Von Eckardt, B. & Potter, M. C. (1985). Clauses and the semantic representation of words. *Memory & Cognition*, *13*, 371–376.
- Walker, I. & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1256–1271.
- Watkins, M. J. (1977). The intricacy of memory span. *Memory & Cognition*, *5*, 529–534.
- Wicha, N. Y. Y., Orozco-Figueroa, A., Reyes, I., Hernandez, A., Gavaldón de Barreto, L., & Bates, E. A. (2005). When zebras become painted donkeys: Grammatical gender and semantic priming interact during picture integration in a spoken Spanish sentence. *Language and Cognitive Processes*, *20*, 553–587.
- Zwitserslood, P. & Bölte, J. (2002). Worterkennung und –produktion. In J. Müsseler, W. Prinz (Eds.), *Allgemeine Psychologie* (pp. 547–587). Heidelberg: Spektrum Akademischer Verlag.

Appendix

A Materials

A 1 German originals

A 1.1 Materials used in Experiment 1, 2, and 3: German originals. Sentences and gender congruent and gender incongruent lure words (in brackets)

1. Nur beim Friseur traute sie sich, das *Magazin*_[neut] (*Heft*_[neut], *Zeitschrift*_[fem]) mit den vielen Klatschgeschichten zu lesen.
2. Der Fahrer stellte das *Fahrzeug*_[neut] (*Auto*_[neut], *Wagen*_[masc]) an diesem Tag ausnahmsweise hinter dem Hotel ab.
3. Die Kollegen hatten ihn mehrfach aufgefordert, sie abends in das *Café*_[neut] (*Restaurant*_[neut], *Kneipe*_[fem]) am Stadtrand zu begleiten.
4. Trotz aller Bemühungen fiel dem Schornsteinfeger die *Klausel*_[fem] (*Regel*_[fem], *Paragraph*_[masc]) für die neue Sicherheitsbestimmung nicht ein.
5. Wie jedes Mal freuten sich besonders die Kinder auf den *Jahrmarkt*_[masc] (*Rummel*_[masc], *Kirmes*_[fem]) mit den vielen Karussells und Süßigkeiten.
6. Der Professor hatte die Änderungen für seine Vorlesung auf ein *Papier*_[neut] (*Blatt*_[neut], *Zettel*_[masc]) geschrieben und suchte nun fieberhaft danach.
7. Er hatte seine Hand schon an den *Knauf*_[masc] (*Griff*_[masc], *Klinke*_[fem]) gelegt, als sich die Tür plötzlich öffnete.
8. Den Detektiv beschlich eine *Vermutung*_[fem] (*Ahnung*_[fem], *Verdacht*_[masc]), was hinter den mysteriösen Vorgängen stecken könnte.
9. Im Anschluss an die *Party*_[fem] (*Feier*_[fem], *Fest*_[neut]) fuhr die ganze Verwandtschaft wieder nach Hause.
10. Nach dem letzten verlustreichen Angriff war die *Truppe*_[fem] (*Armee*_[fem], *Heer*_[neut]) zunächst nicht wieder in die Offensive gegangen
11. Je mehr sie über den Mann hörte, desto größer wurde die *Unsicherheit*_[fem] (*Skepsis*_[fem], *Zweifel*_[masc]) gegenüber seiner Glaubwürdigkeit.
12. Um die Zeitung lesen zu können, war die *Lampe*_[fem] (*Beleuchtung*_[fem], *Licht*_[neut]) in der Küche nicht ausreichend.
13. Der Ganove besitzt eine *Tätowierung*_[fem] (*Narbe*_[fem], *Muttermal*_[neut]), wodurch er besonders leicht zu erkennen ist.
14. Die aufmüpfige Klasse hatte den *Zorn*_[masc] (*Ärger*_[masc], *Wut*_[fem]) der schüchternen Referendarin über die Streiche erheblich unterschätzt.

15. Der vielbeschäftigte Präsident hatte die *Ansprache*_[fem] (*Rede*_[fem], *Vortrag*_[masc]) für die heutige Versammlung nur minimal vorbereitet.
16. Die Kinder betraten aufgeregt den *Saal*_[masc] (*Raum*_[masc], *Zimmer*_[neut]), um endlich die Geschenke auszupacken.
17. Dem genervten Pförtner war die *Freude*_[fem] (*Lust*_[fem], *Spaß*_[masc]) an seinem Beruf durch die arroganten Gäste endgültig verdorben.
18. Der Unternehmer hatte sich noch nicht von dem *Ruin*_[masc] (*Konkurs*_[masc], *Pleite*_[fem]) seiner mühsam aufgebauten Firma erholt.
19. Der Werbetexter wartete auf das *Bild*_[neut] (*Photo*_[neut], *Grafik*_[fem]) für die groß angelegte Kampagne für ein Duschgel.
20. Der Buchhalter hatte einen Termin bei seinem Chef, um über den *Verdienst*_[masc] (*Lohn*_[masc], *Gehalt*_[neut]) neu zu verhandeln.
21. Die Angehörigen stritten heftig um das *Vermächtnis*_[neut] (*Erbe*_[neut], *Erbschaft*_[fem]) des vor drei Tagen verstorbenen Patriarchen.
22. Der Aufseher konnte die *Pein*_[fem] (*Qual*_[fem], *Leid*_[neut]) des Gefangenen nach dem stundenlangen Verhör nicht mehr mit ansehen.
23. Dem vielversprechenden Juraabsolventen war die *Arbeit*_[fem] (*Stelle*_[fem], *Job*_[masc]) in der renommierten Kanzlei schon vor seinem Examen sicher.
24. Sie betraten voller Freude den *Rasen*_[masc] (*Garten*_[masc], *Wiese*_[fem]), um Gänseblümchen und Klee zu pflücken.
25. Wie in jedem Urlaub schrieb Anja die *Post*_[fem] (*Karte*_[fem], *Brief*_[masc]) an ihre Patentante zuerst.
26. Der aufstrebende Jungmanager hatte in der Pause das *Sakko*_[neut] (*Jackett*_[neut], *Jacke*_[fem]) locker über die Schulter geworfen.
27. Vor den Klausuren musste der Schüler immer dringend das *Bad*_[neut] (*Klo*_[neut], *Toilette*_[fem]) in der entsprechenden Etage aufsuchen.
28. Nach der Gartenarbeit war die *Liege*_[fem] (*Couch*_[fem], *Sofa*_[neut]) im Wohnzimmer der einzig mögliche Aufenthaltsort für die Rentnerin.
29. Der Verletzte wurde in aller Eile in das *Hospital*_[neut] (*Krankenhaus*_[neut], *Klinik*_[fem]) gefahren und dort sofort operiert.
30. Der stolze Trainer hatte dem Konkurrenten die *Fähigkeit*_[fem] (*Leistung*_[fem], *Können*_[neut]) seiner kleinen Schützlinge zufrieden demonstriert.

A 1.2 Materials used in Experiment 4: German originals. Sentences and gender congruent and gender incongruent lure words (in brackets) in two versions. In version (a), the target determiner is ambiguously specified for gender, in version (b), it is specified unambiguously.

- 1a. Der Fahrer stand neben *dem Fahrzeug*_[neut] (Auto_[neut], Wagen_[masc]), während er auf seinen Chef wartete.
- 1b. Der Fahrer passte auf *das Fahrzeug*_[neut] (Auto_[neut], Wagen_[masc]) auf, während er auf seinen Chef wartete.
- 2a. Die Änderungen für seine Vorlesung standen auf *einem Papier*_[neut] (Blatt_[neut], Zettel_[masc]), nach dem er nun fieberhaft suchte.
- 2b. Die Änderungen für seine Vorlesung waren auf *ein Papier*_[neut] (Blatt_[neut], Zettel_[masc]) geschrieben, nach dem er nun fieberhaft suchte.
- 3a. Die Angehörigen stritten heftig wegen *des Vermächnisses*_[neut] (Erbe_[neut], Nachlass_[masc]) des vor drei Tagen verstorbenen Patriarchen.
- 3b. Die Angehörigen stritten heftig um *das Vermächtnis*_[neut] (Erbe_[neut], Nachlass_[masc]) des vor drei Tagen verstorbenen Patriarchen.
- 4a. Bezüglich *des Delikts*_[neut] (Verbrechen_[neut], Mord_[masc]) hatte die Polizei nach langer Untersuchung endlich eine Spur.
- 4b. In Bezug auf *das Delikt*_[neut] (Verbrechen_[neut], Mord_[masc]) hatte die Polizei nach langer Untersuchung endlich eine Spur.
- 5a. Die Wissenschaftler wollten mit *einem Test*_[masc] (Versuch_[masc], Experiment_[neut]) mehr über das Gedächtnis erfahren.
- 5b. Die Wissenschaftler wollten durch *den Test*_[masc] (Versuch_[masc], Experiment_[neut]) mehr über das Gedächtnis erfahren.
- 6a. Der Manager verhandelte mit *einem rentablen Konzern*_[masc] (Betrieb_[masc], Unternehmen_[neut]), um Anteile davon aufzukaufen.
- 6b. Der Manager informierte sich über *einen rentablen Konzern*_[masc] (Betrieb_[masc], Unternehmen_[neut]), um Anteile davon aufzukaufen.
- 7a. Sie erbleichte vor *Grausen*_[neut] (Entsetzen_[neut], Schreck_[masc]), als sie das Heulen von draußen hörte.
- 7b. Ihr war *das Grausen*_[neut] (Entsetzen_[neut], Schreck_[masc]) anzusehen, als sie das Heulen von draußen hörte.
- 8a. Die Kinder halfen dabei, die Kartoffeln *vom Land*_[neut] (Feld_[neut], Acker_[masc]) aufzusammeln und in die Körbe zu legen.

- 8b. Die Kinder liefen über *das Land*_[neut] (Feld_[neut], Acker_[masc]) und sammelten die Kartoffeln auf.
- 9a. Sie blieb bis *zum Finale*_[neut] (Ende_[neut], Schluss_[masc]), obwohl ihr das Theaterstück überhaupt nicht gefiel.
- 9b. Sie wartete *das Finale*_[neut] (Ende_[neut], Schluss_[masc]) ab, obwohl ihr das Theaterstück überhaupt nicht gefiel.
- 10a. Der Vater ging mit *seinem Nachwuchs*_[masc] (Sohn_[masc], Kind_[neut]) jeden Sonntag ins Stadion zum Fußballspiel.
- 10b. Der Vater nahm *seinen Nachwuchs*_[masc] (Sohn_[masc], Kind_[neut]) jeden Sonntag ins Stadion zum Fußballspiel mit.
- 11a. Der Verteidigungsminister war extra erschienen, um *dem Schwur*_[masc] (Eid_[masc], Gelöbniß_[neut]) der Soldaten beiwohnen zu können.
- 11b. Der Verteidigungsminister war extra erschienen, um den Soldaten *den Schwur*_[masc] (Eid_[masc], Gelöbniß_[neut]) abnehmen zu können.
- 12a. Die Jugendlichen in *dem kleinen Kaff*_[neut] (Dorf_[neut], Ort_[masc]) trafen sich jeden Abend am Brunnen.
- 12b. Die Jugendlichen liefen durch *das Kaff*_[neut] (Dorf_[neut], Ort_[masc]), um sich am Brunnen zu treffen.
- 13a. Nach zehn Jahren *im Zuchthaus*_[neut] (Gefängnis_[neut], Knast_[masc]) wurde der Häftling wegen guter Führung entlassen.
- 13b. Nach zehn Jahren durfte der Häftling *das Zuchthaus*_[neut] (Gefängnis_[neut], Knast_[masc]) wegen guter Führung verlassen.
- 14a. Der Präsident hatte mit *dem Widerstand*_[masc] (Protest_[masc], Veto_[neut]) gegen das Embargo die anderen Nationen gegen sich aufgebracht.
- 14b. Der Präsident hatte durch *den Widerstand*_[masc] (Protest_[masc], Veto_[neut]) gegen das Embargo die anderen Nationen gegen sich aufgebracht.
- 15a. Die Frau sah aus, als wäre sie *einem Phantom*_[neut] (Gespenst_[neut], Geist_[masc]) begegnet und lief davon.
- 15b. Die Frau sah aus, als hätte sie *ein Phantom*_[neut] (Gespenst_[neut], Geist_[masc]) gesehen und lief davon.
- 16a. Mittags gibt sie sich meist mit *einem kleinen Gericht*_[neut] (Essen_[neut], Imbiss_[masc]) und einer Tasse Kaffee zufrieden.
- 16b. Mittags nimmt sie meist nur *ein kleines Gericht*_[neut] (Essen_[neut], Imbiss_[masc]) und eine Tasse Kaffee zu sich.

- 17a. Nach *seinem Sterben*_[neut] (Ableben_[neut], Tod_[masc]) möchte der Kapitän verbrannt und über dem Atlantik verstreut werden.
- 17b. Im Hinblick auf *sein Sterben*_[neut] (Ableben_[neut], Tod_[masc]) wünscht der Kapitän, verbrannt und über dem Atlantik verstreut zu werden.
- 18a. Die kleine Lena macht immer *viel Lärm*_[masc] (Krach_[masc], Geschrei_[neut]), wenn sie gebadet werden soll.
- 18b. Die kleine Lena macht immer *einen großen Lärm*_[masc] (Krach_[masc], Geschrei_[neut]), wenn sie gebadet werden soll.
- 19a. In Griechenland war sie auf *einem Maultier*_[neut] (Pferd_[neut], Esel_[masc]) geritten und dabei nicht runtergefallen.
- 19b. In Griechenland hatte sie sich auf *ein Maultier*_[neut] (Pferd_[neut], Esel_[masc]) gesetzt und war nicht runtergefallen.
- 20a. Mit *einem vierstimmigen Choral*_[masc] (Kanon_[masc], Stück_[neut]) begannen sie den Gottesdienst in der neuen Kirche.
- 20b. Sie sangen *einen vierstimmigen Choral*_[masc] (Kanon_[masc], Stück_[neut]) und begannen so den Gottesdienst in der neuen Kirche.
- 21a. Der Kommissar suchte nach *einem Anzeichen*_[neut] (Indiz_[neut], Beweis_[masc]), um die Schuld des Mannes festzustellen.
- 21b. Der Kommissar suchte *ein Anzeichen*_[neut] (Indiz_[neut], Beweis_[masc]), um die Schuld des Mannes festzustellen.
- 22a. Während sie im Arbeitsamt wartete, trug sie in *einem Vordruck*_[masc] (Bogen_[masc], Formular_[neut]) schon mal ihre Daten ein.
- 22b. Während sie im Arbeitsamt wartete, trug sie in *einen Vordruck*_[masc] (Bogen_[masc], Formular_[neut]) schon mal ihre Daten ein.
- 23a. Die Vertreter der Röhrenfirma sahen *dem Gespräch*_[neut] (Treffen_[neut], Termin_[masc]) mit den Chinesen freudig entgegen.
- 23b. Die Vertreter der Röhrenfirma freuten sich auf *das Gespräch*_[neut] (Treffen_[neut], Termin_[masc]) mit den Chinesen.
- 24a. Wie Picasso malte er die Seitenansicht *eines Antlitzes*_[neut] (Gesicht_[neut], Kopf_[masc]) mit eckigen Ohren.
- 24b. Wie Picasso malte er das *Antlitz*_[neut] (Gesicht_[neut], Kopf_[masc]) von der Seite mit eckigen Ohren.
- 25a. Sie arbeitete sehr konzentriert an *ihrem Entwurf*_[masc] (Bericht_[masc], Manuskript_[neut]), um dem Gutachter ein Exemplar zu schicken.

- 25b. Sie schrieb sehr konzentriert *ihren Entwurf*_[masc] (Bericht_[masc], Manuskript_[neut]), um dem Gutachter ein Exemplar zu schicken.
- 26a. Die Feuerwehr kämpfte die ganze Nacht mit *einem Flammenmeer*_[neut] (Feuer_[neut], Brand_[masc]) in der Innenstadt.
- 26b. Die Feuerwehr bekämpfte die ganze Nacht *ein Flammenmeer*_[neut] (Feuer_[neut], Brand_[masc]) in der Innenstadt.
- 27a. Kaum jemand merkte, dass nicht Michael Jackson, sondern *ein Imitator*_[masc] (Doppelgänger_[masc], Double_[neut]) auf der Bühne stand.
- 27b. Kaum jemand merkte, dass nicht Michael Jackson, sondern *der Imitator*_[masc] (Doppelgänger_[masc], Double_[neut]) auf der Bühne stand.
- 28a. *Am Horizont*_[masc] (Himmel_[masc], Firmament_[neut]) waren Sterne aufgezogen und es war Vollmond.
- 28b. *Der Horizont*_[masc] (Himmel_[masc], Firmament_[neut]) war voller Sterne und es war Vollmond.
- 29a. Für das Schwein war es *ein Spaß*_[masc] (Genuss_[masc], Vergnügen_[neut]), sich im Schlamm zu suhlen.
- 29b. Für das Schwein war es *der größte Spaß*_[masc] (Genuss_[masc], Vergnügen_[neut]), sich im Schlamm zu suhlen.
- 30a. Auf der Butterfahrt machten sie an *einem urig aussehenden Lokal*_[neut] (Gasthaus_[neut], Gasthof_[masc]) Halt und aßen etwas.
- 30b. Auf der Butterfahrt besuchten sie *ein urig aussehendes Lokal*_[neut] (Gasthaus_[neut], Gasthof_[masc]) und aßen etwas.

A 1.3 Materials used in Experiment 5: German originals. For each item, the context passage and the target region is given. In version (a), the target clause belongs to the penultimate sentence and in version (b), the target clause belongs to the ultimate sentence.

1. Denise hatte sich für den Samstag fest vorgenommen, einen Stadtbummel zu machen. Am frühen Morgen schon ließ sie sich mit ihrem Auto_[neut]/Wagen_[masc]/Hund abholen und nahe der Fußgängerzone absetzen.
 - (a) Der Parkplatz füllte sich langsam, während der Chauffeur das Fahrzeug_[neut] bewachte. Denise ging zum Einkaufen in die Innenstadt.
 - (b) Der Parkplatz füllte sich langsam.

- Während der Chauffeur das Fahrzeug_[neut] bewachte,
ging Denise zum Einkaufen in die Innenstadt.
2. Matze ging auf einen Kaffee bei Bianca vorbei, bevor er an die Uni fuhr. Als er schließlich aufbrach, war er spät dran. Zu allem Überfluss blieb er mit seinem Ärmel an dem Türgriff_[masc]/der Türklinke_[fem]/einem Garderobenhaken hängen.
 - (a) Bianca rief ihn zurück,
als er seine Hand gerade an den Knauf_[masc] legte.
Ihm fiel etwas Wichtiges ein.
 - (b) Bianca rief ihn zurück.
Als er seine Hand gerade an den Knauf_[masc] legte,
fiel ihm etwas Wichtiges ein.
 3. Obwohl der Überfall schon Jahre zurück lag, liefen die Ermittlungen weiter. Lange Zeit hatte niemand eine Ahnung_[fem]/einen Verdacht_[masc]/eine Spur gehabt.
 - (a) Der Täter hatte keine Angst,
obwohl den Detektiv eine Vermutung_[fem] beschlich.
Es war für einen Prozess zu spät.
 - (b) Der Täter hatte keine Angst.
Obwohl den Detektiv eine Vermutung_[fem] beschlich,
war es für einen Prozess zu spät.
 4. Am Weihnachtsabend aß die Familie abends Braten und Kartoffelklöße. Während die Kinder vor der Bescherung noch aufräumen mussten, schmückte der Vater den Raum_[masc]/das Zimmer_[neut]/den Weihnachtsbaum.
 - (a) Die Kinder freuten sich,
als sie aufgeregt den Saal_[masc] betraten.
Sie durften die Geschenke auspacken.
 - (b) Die Kinder freuten sich.
Als sie aufgeregt den Saal_[masc] betraten,
durften sie die Geschenke auspacken.
 5. Bis vor kurzem war Dieter Besitzer einer kleinen Firma für Bodenverlegung. Es sah sogar noch ganz gut aus, als die Hauptpost ihn beauftragte, 500 m² PVC zu verlegen. Doch als dann ein Zulieferer Konkurs_[masc]/Pleite_[fem]/Schwierigkeiten machte, konnte er den Auftrag nicht mehr erfüllen.
 - (a) Jetzt braucht Dieter viel Ruhe,
um den Ruin_[masc] seiner Firma zu verkraften.

- Er macht Urlaub in der Karibik.
- (b) Jetzt braucht Dieter viel Ruhe.
Um den Ruin_[masc] seiner Firma zu verkraften,
macht er Urlaub in der Karibik.
6. Der Leiter der kleinen Werbeagentur erhielt einen unangenehmen Anruf. Die Auftraggeber für eine wichtige Kampagne waren nicht bereit, Verzögerungen hinzunehmen. Verlegen sah er die Bilder_[neut]/Grafiken_[fem]/Unterlagen auf seinem Schreibtisch durch.
- (a) Die Atmosphäre in der Agentur war gespannt,
weil der Werbetexter auf die Fotos_[neut] wartete.
Die Plakate waren noch nicht fertig.
- (b) Die Atmosphäre in der Agentur war gespannt.
Weil der Werbetexter auf die Fotos_[neut] wartete,
waren die Plakate noch nicht fertig.
7. Der Buchhalter war sich darüber im Klaren, dass die Unternehmen zurzeit nicht viel Geld für ihre Angestellten hatten. Dennoch musste er es versuchen. Letzten Monat war ihm die Miete erhöht worden, und er musste umziehen, wenn er nicht mehr Lohn_[masc]/Gehalt_[neut]/Zulagen bekäme.
- (a) Der Buchhalter hatte einen Termin bei seinem Chef,
um über seinen Verdienst_[masc] zu verhandeln.
Er bat um die Unterstützung der Gewerkschaft.
- (b) Der Buchhalter hatte einen Termin bei seinem Chef.
Um über seinen Verdienst_[masc] zu verhandeln,
bat er um die Unterstützung der Gewerkschaft.
8. Anja hatte zum Geburtstag von ihrer Patentante einige tolle Geschenke bekommen. Heute wollte sie sich nun angemessen bedanken. Sie setzte sich an ihren Schreibtisch und begann die Karte_[fem]/den Brief_[masc]/mit einer herzlichen Begrüßung.
- (a) Die Sonne schien durchs Fenster,
während Anja die Post_[fem] an ihre Patentante schrieb.
Ihre Mutter lag gemütlich auf der Terrasse.
- (b) Die Sonne schien durchs Fenster.
Während Anja die Post_[fem] an ihre Patentante schrieb,
lag ihre Mutter gemütlich auf der Terrasse.

9. Die Schüler gingen den Gang vor dem Prüfungssaal unruhig auf und ab und rauchten dabei eine Zigarette nach der anderen. Viele waren so auf die bevorstehende Abschlussarbeit konzentriert, dass sie kaum die Putzfrauen wahrnahmen, die das Klo_[neut]/die Toilette_[fem]/den Flur sauber machten.
- (a) Stefan war nervös und aufgeregt,
weil das Bad_[neut] vor der Klausur besetzt war.
Er traute sich nicht mehr zu trinken.
- (b) Stefan war nervös und aufgeregt.
Weil das Bad_[neut] vor der Klausur besetzt war,
traute er sich nicht mehr zu trinken.
10. Nach langjähriger Krankheit verstarb sein Vater im Krankenhaus. Die Familie litt sehr unter dem Verlust. Besonders der Sohn trauerte, da er seinen Vater seit fünf Jahren nicht mehr gesehen hatte. Nun saß die Familie zusammen, um über das Erbe_[neut]/die Erbschaft_[fem]/das Begräbnis des Vaters zu sprechen.
- (a) Der Notar war gerade erschienen,
weil das Vermächtnis_[neut] des Familienoberhaupts unklar war.
Die Angehörigen stritten sich heftig.
- (b) Der Notar war gerade erschienen.
Weil das Vermächtnis_[neut] des Familienoberhaupts unklar war,
stritten die Angehörigen sich heftig.
11. Berta musste eigentlich noch ein Kapitel „Faust II“ für den Deutschunterricht am nächsten Tag lesen. Sie mochte das Buch nicht besonders und hätte lieber etwas mit einer Freundin unternommen. Sie riss sich zusammen, suchte sich einen ruhigen Ort in der Wohnung und machte die Lampe_[fem]/das Licht_[neut]/den Ventilator an.
- (a) Das Lesen machte keinen Spaß,
weil die Beleuchtung_[fem] in der Küche nicht ausreichte.
Berta ging in ein anderes Zimmer.
- (b) Das Lesen machte keinen Spaß.
Weil die Beleuchtung_[fem] in der Küche nicht ausreichte,
ging Berta in ein anderes Zimmer.
12. Die Wissenschaftler untersuchen schon seit einiger Zeit das Gedächtnis. Die neuesten Forschungsergebnisse waren ein großer Erfolg. Um noch weitere Er-

- kenntnisse zu erlangen wurde ein neuer Versuch_[masc]/ein neues Experiment_[neut]/
eine neue Vorgehensweise geplant.
- (a) Viel Vorbereitung war erforderlich,
um den Test_[masc] durchzuführen.
Mehrere Hilfskräfte wurden eingestellt.
- (b) Viel Vorbereitung war erforderlich.
Um den Test_[masc] durchzuführen,
wurden mehrere Hilfskräfte eingestellt.
13. Auf einer von vielen Pendlern benutzten Zugstrecke kam es wegen Bauarbeiten
zu einem falschen Signal. Zwei voll besetzte Züge kollidierten. Zum Glück wa-
ren beide nicht schnell gefahren, dennoch gab es viele Verletzte. Außerdem hat-
ten die Zugbegleiter zunächst technische Schwierigkeiten, das Krankenhaus_[neut]/
die Klinik_[fem]/den Notarzt zu benachrichtigen.
- (a) Die Angehörigen konnten nur bangen,
während die Verletzten in das Hospital_[neut] gefahren wurden.
Die Ärzte bereiteten den Operationssaal vor.
- (b) Die Angehörigen konnten nur bangen.
Während die Verletzten in das Hospital_[neut] gefahren wurden,
bereiteten die Ärzte den Operationssaal vor.
14. Nachdem sie nur drei Monate an der neuen Schule unterrichtet hatte, bekam die
Lehrerin überraschend Besuch von der Schulaufsichtsbehörde. Es sei nur eine
Routine-Lehrprobe, um zu sehen ob alles in Ordnung ist. Die Lehrerin ahnte
schon, dass die unerwartete Störung des Unterrichts durch die drei Männer im
Anzug ihre Schüler irritieren würde. Zuerst hatte sie ausgerechnet die achte Klas-
se, und einige der pubertierenden Jugendlichen nutzten die Gelegenheit, um die
Lehrerin lächerlich zu machen. Diese Clique hatte ihr schon einige Male Är-
ger_[masc]/Wut_[fem]/Bauchweh bereitet.
- (a) Die Lehrerin musste sich sehr anstrengen,
um ihren Zorn_[masc] zu unterdrücken.
Sie atmete mehrmals kräftig durch.
- (b) Die Lehrerin musste sich sehr anstrengen.
Um ihren Zorn_[masc] zu unterdrücken,
atmete sie mehrmals kräftig durch.

15. Die Jahresvollversammlung des Philatelieverbandes stand unter keinem guten Stern. Das Kongresszentrum, in dem sie sonst immer einen Saal gemietet hatten, war dieses Jahr bereits ausgebucht gewesen. Eigentlich hatte keines der Vorstandsmitglieder genug Zeit um alles zu organisieren. Nun fand das Treffen in dem Hinterzimmer eines zweitklassigen Hotels statt. Schon vor Beginn des Programms meinten mehrere der anwesenden Mitglieder, dass sie sich die Rede_[fem]/den Vortrag_[masc]/das Essen wohl besser erspart hätten.
- (a) Der Präsident musste heute viel improvisieren, weil er seine Ansprache_[fem] kaum vorbereitet hatte. Die Teilnehmer hörten ihm überhaupt nicht zu.
- (b) Der Präsident musste heute viel improvisieren. Weil er seine Ansprache_[fem] kaum vorbereitet hatte, hörten ihm die Teilnehmer überhaupt nicht zu.
16. Marcel war etwas mulmig, als er unter Hunderten anderer Rekruten der Bundeswehr bewegungslos dastand. Zum einen hatte er schlecht geschlafen, zum anderen nicht gefrühstückt. Beides verstärkte noch seine Unruhe und Nervosität. Er hätte nie gedacht, dass er sich einmal dazu durchringen würde, einen Eid_[masc]/ein Gelöbnis_[neut]/Wehrdienst zu leisten.
- (a) Der Verteidigungsminister war extra erschienen, um den Soldaten den Schwur_[masc] abzunehmen. Alle hatten sich auf dem großen Platz versammelt.
- (b) Der Verteidigungsminister war extra erschienen. Um den Soldaten den Schwur_[masc] abzunehmen, hatten sich alle auf dem großen Platz versammelt.
17. Nach Jahren der Abwesenheit kehrte sie in ihre Heimat zurück. Es fiel ihr nicht ganz leicht; sie hatte sich verändert, und sie fürchtete auch die Veränderungen der kleinen Stadt. Ob sie überhaupt noch etwas wieder erkennen würde? Um sich das Vorhaben zu versüßen, hatte sie geplant, sich eine erinnerungsträchtige Süßigkeit zu gönnen. Nun jedoch musste sie feststellen, dass ein Einkaufszentrum_[neut]/ein Supermarkt_[masc]/ein Kundenparkplatz den Platz ihres Lieblingsbäckers eingenommen hatte.
- (a) Das Verkehrsaufkommen hatte deutlich zugenommen, seit das Geschäft_[neut] neu eröffnet worden war. Immer weniger Kunden gingen in die kleinen Läden.

- (b) Das Verkehrsaufkommen hatte deutlich zugenommen.
Seit das Geschäft_[neut] neu eröffnet worden war,
gingen immer weniger Kunden in die kleinen Läden.
18. Martina hatte ihre gesamte Schulzeit nur mit dem Ziel hinter sich gebracht,
einmal Tiermedizin zu studieren. Nun war es soweit, sie hatte es geschafft.
Gleich in der ersten Semesterwoche besuchte sie einen Präparationskurs. Als sie
den Saal betrat und die Tierleichen überall auf den Tischen liegen sah, umgeben
von blitzenden Instrumenten und Werkzeugen, überkam sie unvermittelt Entset-
zen_[neut]/Angst_[fem]/Übelkeit.
- (a) Sie wollte dableiben und durchhalten,
obwohl ihr das Grauen_[neut] ins Gesicht geschrieben stand.
Sie wandte den Blick nicht ab.
- (b) Sie wollte dableiben und durchhalten.
Obwohl ihr das Grauen_[neut] ins Gesicht geschrieben stand,
wandte sie den Blick nicht ab.
19. Die Bundestagswahlen standen vor der Tür. Jana fand das nicht besonders inte-
ressant, da sie ohnehin noch nicht wählen durfte. Trotzdem wurde ihre Klasse
von der Politik-Lehrerin dauernd gezwungen, Sendungen_[fem]/Berichte_[masc]/Infor-
mationen darüber zu sammeln und in den Unterricht einzubringen.
- (a) Jana schimpfte über ihre Lehrerin,
weil sie sich die Reportage_[fem] über die Wahl ansehen musste.
Sie verpasste heute ihre Lieblingsserie.
- (b) Jana schimpfte über ihre Lehrerin.
Weil sie sich die Reportage_[fem] über die Wahl ansehen musste,
verpasste sie heute ihre Lieblingsserie.
20. Schon seit geraumer Zeit machte ein Nachbar Frau Fischer den Hof. Sie war im
Grunde nicht uninteressiert, ließ es aber ganz langsam angehen. Kürzlich war sie
mit ihm ins Theater gegangen. Zunächst hatte sie sich geziert, die Einladung an-
zunehmen, aber am Ende_[neut]/zum Schluss_[masc]/zu guter letzt hatte sie ihrer ersten
Verabredung zugestimmt.
- (a) Das Theaterstück hatte ihr überhaupt nicht gefallen,
obwohl sie bis zum Finale_[neut] geblieben war.
Sie hatte nicht verstanden, worum es ging.
- (b) Das Theaterstück hatte ihr überhaupt nicht gefallen.

- Obwohl sie bis zum Finale_[neut] geblieben war,
hatte sie nicht verstanden, worum es ging.
21. Eine stadtbekannte Motorradgang machte immer wieder durch typische Gewalttaten auf sich aufmerksam. Die Bande suchte sich einzelne Opfer aus und ließ nicht von ihnen ab, bis sie blutend im Straßengraben lagen. Nun waren sie sogar soweit gegangen, ihr Opfer zu erstechen. Zumindest war man sich ziemlich sicher, dass dieses Verbrechen_[neut]/diese Tat_[fem]/diese Grausamkeit auch auf ihr Konto ging.
- (a) Die Polizei suchte nach einem Zeugen,
um der Bande das Delikt_[neut] nachzuweisen.
Bisher hatten sich nur Indizien gefunden.
- (b) Die Polizei suchte nach einem Zeugen.
Um der Bande das Delikt_[neut] nachzuweisen.
hatten sich bisher nur Indizien gefunden.
22. Die Medien hatten gejubelt, als sein Trainervertrag unter Dach und Fach war. Auch der Trainer selbst hatte sich auf die Aufgabe gefreut. Viele der Spieler waren jung und motiviert. Doch in letzter Zeit hatten sie viele Fehlschläge_[masc]/Niederlagen_[fem]/Gegentore einstecken müssen.
- (a) Der Trainer wurde immer unruhiger,
weil sein Misserfolg_[masc] die Mannschaft auf einen Abstiegsplatz gebracht hatte.
Er fürchtete um seinen gut bezahlten Job.
- (b) Der Trainer wurde immer unruhiger.
Weil sein Misserfolg_[masc] die Mannschaft auf einen Abstiegsplatz gebracht hatte,
fürchtete er um seinen gut bezahlten Job.
23. Der Prozess war spektakulär. Bis zur Verhandlung wurde in den Nachrichten erörtert, wie das Gericht wohl entscheiden würde. Es war klar, dass der Angeklagte sein ganzes Leben unter extremen Bedingungen gelitten hatte. Da er sich aber schließlich des Mordes schuldig gemacht hatte, blieb ihm der Knast_[masc]/das Gefängnis_[neut]/ein Schuldspruch nicht erspart.
- (a) Der Angeklagte erregte bei vielen Zuschauern Mitleid,
weil er lebenslänglich in den Bau_[masc] geschickt wurde.
Er brach noch bei der Verhandlung in Tränen aus.
- (b) Der Angeklagte erregte bei vielen Zuschauern Mitleid.
Weil er lebenslänglich in den Bau_[masc] geschickt wurde,

brach er noch bei der Verhandlung in Tränen aus.

24. Der Sicherheitsrat sah die Möglichkeit, durch ein Embargo die Diktatur in dem kleinen Land zu beenden. Fast alle Mitglieder waren darin einer Meinung. Der Vorsitzende verhandelte jedoch vergeblich mit einem Präsidenten, der den Beschluss mit seinem Widerspruch_[masc]/Veto_[neut]/Starrsinn kippen wollte.
- (a) Der Präsident machte sich im Sicherheitsrat Feinde, als er seinem Protest_[masc] gegen das Embargo aussprach. Er brachte die anderen Nationen damit gegen sich auf.
- (b) Der Präsident machte sich im Sicherheitsrat Feinde. Als er seinem Protest_[masc] gegen das Embargo aussprach, brachte er die anderen Nationen damit gegen sich auf.
25. Die Region war besonders von Arbeitslosigkeit und Landflucht betroffen. Seit das Kino aus finanziellen Gründen geschlossen worden war, hatte das Dorf_[neut]/der Ort_[masc]/die Gegend seine letzte Attraktion verloren.
- (a) Die Jugendlichen trafen sich jetzt jeden Abend am Brunnen, weil sie in dem Kaff_[neut] nichts anderes tun konnten. Ihnen war oft langweilig.
- (b) Die Jugendlichen trafen sich jetzt jeden Abend am Brunnen. Weil sie in dem Kaff_[neut] nichts anderes tun konnten, war ihnen oft langweilig.
26. Detlev war mittlerweile Mitte 20 und studierte nicht weit von der Stadt, in der er aufgewachsen war. Er war ein ruhiges und freundliches Kind gewesen, das seinen Eltern niemals Schwierigkeiten_[fem]/Probleme_[neut]/Kummer bereitete.
- (a) Detlev besuchte noch oft seine Mutter, um Hilfe bei allen Fragen_[fem] zu bekommen. Er wandte sich jedes Mal an sie.
- (b) Detlev besuchte noch oft seine Mutter. Um Hilfe bei allen Fragen_[fem] zu bekommen, wandte er sich jedes Mal an sie.
27. An ihrem Hochzeitstag wollten der Bauer und die Bäuerin abends fein Essen und anschließend ins Kino gehen. Um rechtzeitig Feierabend machen zu können, mussten sie beide besonders früh aufstehen. Das Feld_[neut]/der Acker_[masc]/neues Futter musste noch dringend bestellt werden.
- (a) Der Bauer fuhr mit seinem Traktor los,

- um das Land_[neut] zu pflügen.
Er musste einen ganzen Nachmittag einplanen.
- (b) Der Bauer fuhr mit seinem Traktor los.
Um das Land_[neut] zu pflügen,
musste er einen ganzen Nachmittag einplanen.
28. Der Gagschreiber arbeitete an einer neuen Show für einen berühmten Komiker, der gerade in einer Krise steckte. Er selbst hatte schon genügend kreative Tiefpunkte durchlebt. Er wusste, wie es sich anfühlt, wenn nichts Lustiges aus der Feder fließt. Dafür wusste er allerdings auch jeden guten Einfall_[masc]/jede gute Idee_[fem]/jeden guten Lacher zu feiern.
- (a) Der Gagschreiber war stolz auf sich,
weil ihm der Gedanke_[masc] für den Sketch so gut gefiel.
Er nahm sich den Rest des Tages frei.
- (b) Der Gagschreiber war stolz auf sich.
Weil ihm der Gedanke_[masc] für den Sketch so gut gefiel,
nahm er sich den Rest des Tages frei.
29. Anna und Kati waren die besten Freundinnen, die man sich vorstellen konnte. Wann immer sie sich trafen spielten und plauderten sie und vergaßen dabei die Zeit. Bei schönem Wetter gingen sie meistens an die frische Luft und spielten im Garten_[masc]/auf der Wiese_[fem]/hinter dem Haus.
- (a) Die beiden Mädchen lachten und hüpfen,
Wenn sie den Rasen_[masc] betreten.
Sie pflückten Gänseblümchen und Klee.
- (b) Die beiden Mädchen lachten und hüpfen.
Wenn sie den Rasen_[masc] betreten,
pflückten sie Gänseblümchen und Klee.
30. Roland hatte eine sehr religiöse Freundin. Für sie war Aberglaube mit Blasphemie gleichzusetzen. Schließlich gerieten sie in erbitterten Streit darüber, ob es mit echtem Gottesglauben vereinbar sei, ein Maskottchen_[neut]/einen Talisman_[masc]/einen Bibelvers zu Vorstellungsgesprächen mitzunehmen.
- (a) Roland hielt sich nicht für abergläubisch,
obwohl er immer sein Hufeisen_[neut] dabei hatte.
Er glaubte weder an Horoskope noch an Wahrsager.
- (b) Roland hielt sich nicht für abergläubisch.

Obwohl er immer sein Hufeisen_[neut] dabei hatte,
glaubte er weder an Horoskope noch an Wahrsager.

A 2 English glosses

Note that the translations for the target word and the two types of lure words can only be approximations. As there is obviously no gender difference for these words in English, the suggested translations only serve to indicate that more than one word is a better fit for the sentence content than the target word (although the English synonym triplets (target word, gender congruent lure word, and gender incongruent lure word) are not equivalent to the German synonym triplets.)

A 2.1 Materials used in Experiment 1, 2, and 3: English glosses. Sentences and gender congruent and gender incongruent lure words (in brackets)

1. Only at the barber's shop she dared reading the newspaper (journal, magazine) full of gossip.
2. That day, the driver exceptionally parked the vehicle (auto, car) at the back of the hotel.
3. The colleagues had often asked him to accompany them to the café (restaurant, pub) on the outskirts.
4. Despite all efforts the chimney sweeper could not remember the rule (norm, paragraph) for the new safety regulation.
5. As always, especially the children were looking forward to the kermes (carnival, fair) with the many carousels and sweets.
6. The changes for the professor's lecture were written on a note (sheet, paper) which he was searching frenetically now.
7. As he put his hand on the latch (knob, handle) the door suddenly opened.
8. An assumption (notion, suspicion) came to the investigator's mind about what might have caused those mysterious happenings.
9. After the festivity (celebration, party) all relatives drove home again.
10. Since the last attack which had involved heavy losses the corps (troops, army) had not yet started another offensive.
11. The more she heard about the guy, the more her scepticism (uncertainty, doubt) about his credibility increased.

12. The illumination (lamp, light) in the kitchen was not strong enough to allow for reading the newspaper.
13. The gangster can easily be identified by his tattoo (scar, mole).
14. The impudent class had substantially underestimated the shy assistant teacher's ire (anger, rage) about their capers.
15. The busy president had hardly prepared the address (talk, speech) for today's assembly.
16. The children who wanted to open the presents eventually excitedly entered the hall (chamber, room).
17. The distressed porter had finally lost delight (pleasure, joy) in his job because of the arrogant guests.
18. The manager had not yet recovered from the insolvency (bankruptcy, crash) of his firm.
19. The ad writer had not yet received the picture (photo, graphic) for the large-scale shower gel campaign.
20. The accountant had a meeting with his boss in order to discuss his wages (salary, income).
21. The relatives were quarrelling seriously about the recently deceased patriarch's devise (legacy, heritage).
22. The prison guard could not stand to see the prisoner's agony (distress, suffering) after the long interrogation.
23. The ambitious law student had been promised a work (employment, job) in the famous law office before his exams already.
24. They happily entered the lawn (garden, meadow) in order to pick daisies and clover.
25. As always in her vacation Anja wrote the mail (postcard, letter) to her godmother first.
26. During the break the ambitious young manager had thrown the blazer (coat, jacket) loosely over his shoulder.
27. Before the exams the pupil always had to go to the lavatory (toilet, bathroom) on the respective floor.
28. After the garden work the divan (couch, sofa) in the living room was the only possible place to be for the retiree.

29. They took the injured one to the infirmary (clinic, hospital) and surgically supplied him immediately.
30. The coach had proudly presented his little protégés abilities (achievements, skills) to his competitor.

A 2.2 Materials used in Experiment 4: English glosses. Sentences and gender congruent and gender incongruent lure words (in brackets) in two versions. In version (a), the target determiner in the German original is ambiguously specified for gender, in version (b), it is specified unambiguously.

- 1a. The driver stood alongside the vehicle (car, automobile) while he was waiting for his boss.
- 1b. The driver watched the vehicle (car, automobile) while he was waiting for his boss.
- 2a. The changes for his lecture were on a paper (sheet, note) which he was searching for desperately now.
- 2b. The changes for his lecture were written on a paper (sheet, note) which he was searching for desperately now.
- 3a. The family had a dispute because of the bequest (inheritance, legacy) from their patriarch who had died three days ago.
- 3b. The family had a dispute about the bequest (inheritance, legacy) from their patriarch who had died three days ago.
- 4a. After a lengthy inquiry, the police finally got a clue concerning the offence (crime, murder).
- 4b. After a lengthy inquiry, the police finally got a clue with regard to the offence (crime, murder).
- 5a. By means of a test (study, experiment), the scientists attempted to find out more about memory.
- 5b. Through a test (study, experiment), the scientists attempted to find out more about memory.
- 6a. The manager negotiated with a profitable business (company, enterprise) in order to buy capital shares of it.
- 6b. The manager gathered information about a profitable business (company, enterprise) in order to buy capital shares of it.
- 7a. She whitened in terror (horror, alarm) when she heard the howling from outside.

- 7b. One could see the terror (horror, alarm) in her face when she heard the howling from outside.
- 8a. The children assisted in picking up the potatoes off the land (ground, acre) and in putting them into baskets.
- 8b. The children walked across the land (ground, acre) and picked up the potatoes.
- 9a. She stayed to the finish (finale, close) although she did not like the play at all.
- 9b. She waited for the finish (finale, close) although she did not like the play at all.
- 10a. On Sundays, the father always went to the stadium with his offspring (son, child) to watch football.
- 10b. On Sundays, the father always took his offspring (son, child) to the stadium to watch football.
- 11a. The Minister of Defence had appeared in order to be present at the soldiers' vow (oath, pledge).
- 11b. The Minister of Defence had appeared in order to administer the vow (oath, pledge) to the soldiers.
- 12a. In the evenings, the teenagers in the small backwater (village, place) met at the well.
- 12b. The teenagers walked through the small backwater (village, place) in order to meet at the well.
- 13a. After ten years, the prisoner was released from penitentiary (jail, prison) because of proper conduct.
- 13b. After ten years, the prisoner was allowed to leave penitentiary (jail, prison) because of proper conduct.
- 14a. With his resistance (opposition, veto) against the embargo the president provoked the hostility of the other nations.
- 14b. Through his resistance (opposition, veto) against the embargo the president provoked the hostility of the other nations.
- 15a. The woman looked as if she had come across a spirit (ghost, phantom) and ran away.
- 15b. The woman looked as if she had seen a spirit (ghost, phantom) and ran away.
- 16a. At noon, she is usually content with a small meal (snack, lunch) and a cup of coffee.
- 16b. At noon, she usually has nothing but a small meal (snack, lunch) and a cup of coffee.

- 17a. After his end (death, decease), the captain would be cremated and scattered over the Atlantic.
- 17b. With regard to his end (death, decease), the captain would be cremated and scattered over the Atlantic.
- 18a. Little Lena always makes a lot of hullabaloo (noise, clamour) when she is to have a bath.
- 18b. Little Lena always makes a great hullabaloo (noise, clamour) when she is to have a bath.
- 19a. In Greece, she had had a ride on a dinky (donkey, horse) without falling down.
- 19b. In Greece, she had bestridden a dinky (donkey, horse) without falling down.
- 20a. They began the service in the new church with a four-part chorale (round, song).
- 20b. They began the service in the new church by singing a four-part chorale (round, song).
- 21a. The police inspector tried to find an indication (proof, evidence) in order to ascertain the man's guilt.
- 21b. The police inspector was searching for an indication (proof, evidence) in order to ascertain the man's guilt.
- 22a. She already wrote her data on a blank (sheet, form) while she was waiting at the employment bureau.
- 22b. She already wrote her data on a blank (sheet, form) while she was waiting at the employment bureau.
- 23a. The representatives of the pipe company were gladly anticipating the colloquy (date, meeting) with the Chinese.
- 23b. The representatives of the pipe company were looking forward to the colloquy (date, meeting) with the Chinese.
- 24a. Just like Picasso he drew the side view of a countenance (face, head) with angular ears.
- 24b. Just like Picasso he drew a countenance (face, head) in side view with angular ears.
- 25a. She worked very hard at her blueprint (article, manuscript) in order to send a copy to her reviewer.
- 25b. She wrote her blueprint (article, manuscript) very concentratedly in order to send a copy to her reviewer.
- 26a. All night long, the firemen fought combustion (conflagration, fire) in the city.

- 26b. All night long, the firemen tried to extinguish combustion (conflagration, fire) in the city.
- 27a. Hardly anyone noticed that it was not Michael Jackson who was on stage but a mimic (double, imitator).
- 27b. Hardly anyone noticed that it was not Michael Jackson who was on stage but the mimic (double, imitator).
- 28a. On the horizon (sky, firmament), stars had raised and the moon was full.
- 28b. The horizon (sky, firmament) was abounding with stars and the moon was full.
- 29a. It was great fun (pleasure, delight) for the pig to wallow in the mud.
- 29b. It was the greatest fun (pleasure, delight) for the pig to wallow in the mud.
- 30a. On a day trip they stopped at a nice looking pub (inn, tavern) and had something to eat.
- 30b. On a day trip they visited a nice looking pub (inn, tavern) and had something to eat.

A 2.3 Materials used in Experiment 5: English glosses. Target region and gender congruent and gender incongruent lure words (in brackets). In version (a), the target clause belongs to the penultimate sentence and in version (b), the target clause belongs to the ultimate sentence.

- 1a. The car park filled up slowly
while the chauffeur was watching the vehicle (auto, car).
Denise went downtown for shopping.
- 1b. The car park filled up slowly.
While the chauffeur was watching the vehicle (auto, car),
Denise went downtown for shopping.
- 2a. Bianca called him back
when he was just touching the latch (knob, handle).
Something important crossed her mind.
- 2b. Bianca called him back.
When he was just touching the latch (knob, handle),
something important crossed her mind.
- 3a. The perpetrator was not afraid,
although a supposition (hint, suspicion) crept into the detective's mind.
It was too late for judicial proceedings.

- 3b. The perpetrator was not afraid.
Although a supposition (hint, suspicion) crept into the detective's mind,
it was too late for judicial proceedings.
- 4a. The children were delighted
when they excitedly entered the hall (chamber, room).
They were allowed to unwrap the presents.
- 4b. The children were delighted.
When they excitedly entered the hall (chamber, room),
they were allowed to unwrap the presents.
- 5a. Dieter needs a long rest now
to get over his company's ruin (bankruptcy, insolvency).
He is on holiday in the Caribbean now.
- 5b. Dieter needs a long rest now.
To get over his company's ruin (bankruptcy, insolvency),
he is on holiday in the Caribbean now.
- 6a. The atmosphere in the company was tense
because the adman was waiting for the photos (pictures, graphics).
The poster wasn't finished yet.
- 6b. The atmosphere in the company was tense.
Because the adman was waiting for the photos (pictures, graphics),
the poster wasn't finished yet.
- 7a. The accountant had an appointment with his boss
in order to bargain for his earnings (salary, income).
He asked the labor union for support.
- 7b. The accountant had an appointment with his boss.
In order to bargain for his earnings (salary, income),
he asked the labor union for support.
- 8a. The sun was shining through the window,
while Anja was writing a mail (postcard, letter) to her godmother.
Her mother cuddled on the terrace.
- 8b. The sun was shining through the window.
While Anja was writing a mail (postcard, letter) to her godmother,
her mother cuddled on the terrace.
- 9a. Stefan was nervous and excited

- because the toilet (restroom, bathroom) was occupied before the exam.
He did not dare to drink anymore.
- 9b. Stefan was nervous and excited.
Because the toilet (restroom, bathroom) was occupied before the exam,
he did not dare to drink anymore.
- 10a. The notary had just arrived
because the head of the family's bequest (heritage, legacy) was unclear.
The relatives disputed bitterly.
- 10b. The notary had just arrived.
Because the head of the family's bequest (heritage, legacy) was unclear,
the relatives disputed bitterly.
- 11a. It was no fun reading
as the illumination (light, lamp) in the kitchen was insufficient.
Berta went to another room.
- 11b. It was no fun reading.
As the illumination (light, lamp) in the kitchen was insufficient,
Berta went to another room.
- 12a. A lot of preparations were necessary
in order to carry out the test (study, experiment).
several assistants were employed.
- 12b. A lot of preparations were necessary.
In order to carry out the test (study, experiment),
several assistants were employed.
- 13a. The relatives could just wait and hope,
while the injured persons were taken to the infirmary (hospital, emergency
room).
The physicians prepared the surgery room.
- 13b. The relatives could just wait and hope.
While the injured persons were taken to the infirmary (hospital, emergency
room),
the physicians prepared the surgery room.
- 14a. The teacher had to struggle hard
in order to suppress her fury (anger, rage).
She took more than one deep breath.

- 14b. The teacher had to struggle hard.
In order to suppress her fury (anger, rage),
she took more than one deep breath.
- 15a. The president had to improvise a lot this day
because he had hardly prepared his address (talk, speech).
The participants didn't listen to him at all.
- 15b. The president had to improvise a lot this day.
Because he had hardly prepared his address (talk, speech),
the participants didn't listen to him at all.
- 16a. The Minister of Defence had appeared
in order to put the soldiers upon their vows (oaths, pledges).
Everyone had gathered in the wide place.
- 16b. The Minister of Defence had appeared.
In order to put the soldiers upon their vows (oaths, pledges),
everyone had gathered in the wide place.
- 17a. There had been a remarkable increase in traffic volume
since the magasin (supermarket, shopping-center) had newly opened.
Less and less customers went to the smaller stores.
- 17b. There had been a remarkable increase in traffic volume
Since the magasin (supermarket, shopping-center) had newly opened,
less and less customers went to the smaller stores.
- 18a. She wanted to stay and to see through,
although terror (horror, fear) could be observed in her face.
She didn't turn away her eyes.
- 18b. She wanted to stay and to see through.
Although terror (horror, fear) could be observed in her face,
she didn't turn away her eyes.
- 19a. Jana railed against her teacher
because she had to watch the report (broadcast, information) about the elections.
She missed her favorite series today.
- 19b. Jana railed against her teacher.
Because she had to watch the report (broadcast, information) about the elections,
she missed her favorite series today.
- 20a. She didn't enjoy the play at all,

- although she had stayed up to the closure (end, finale).
She didn't understand what it was about.
- 20b. She didn't enjoy the play at all.
Although she had stayed up to the closure (end, finale),
she didn't understand what it was about.
- 21a. The police was looking for a witness
in order to trace the felony (crime, offence) back to this gang.
Nothing but evidences had been found until now.
- 21b. The police was looking for a witness.
In order to trace the felony (crime, offence) back to this gang,
nothing but evidences had been found until now.
- 22a. The coach became ever more nervous,
as his washout (failure, defeat) had brought his team on its way down.
He was afraid of losing his well-paid job.
- 22b. The coach became ever more nervous.
As his washout (failure, defeat) had brought his team on its way down,
he was afraid of losing his well-paid job.
- 23a. The culprit aroused sympathy in many of the spectators
because he was sent to penitentiary (prison, jail) for life.
He burst into tears during the trial.
- 23b. The culprit aroused sympathy in many of the spectators.
Because he was sent to penitentiary (prison, jail) for life,
he burst into tears during the trial.
- 24a. The president made enemies in the Security Council
when he spoke out a protest (veto, negative) against the embargo.
He infuriated the other nations.
- 24b. The president made enemies in the Security Council.
When he spoke out a protest (veto, negative) against the embargo,
he infuriated the other nations.
- 25a. The youngsters now met at the well in the evenings
because they couldn't do anything else in this dump (village, town).
They were bored frequently.
- 25b. The youngsters now met at the well in the evenings.
Because they couldn't do anything else in this dump (village, town),

- they were bored frequently.
- 26a. Detlev often saw his mother
in order to get help with any question (difficulty, problem).
He addressed her each time.
- 26b. Detlev often saw his mother.
In order to get help with any question (difficulty, problem),
he addressed her each time.
- 27a. The farmer started with his tractor
in order to plough the land (field, soil).
He had to allow for needing the whole afternoon.
- 27b. The farmer started with his tractor.
In order to plough the land (field, soil),
he had to allow for needing the whole afternoon.
- 28a. The skit writer was very proud of himself
because he really liked the suggestion (idea, inspiration) for the sketch.
He spent the rest of that day leisurely.
- 28b. The skit writer was very proud of himself.
Because he really liked the suggestion (idea, inspiration) for the sketch,
he spent the rest of that day leisurely.
- 29a. The two girls laughed and bounced
as they entered the lawn (garden, meadow).
They collected daisies and clover.
- 29b. The two girls laughed and bounced.
As they entered the lawn (garden, meadow),
they collected daisies and clover.
- 30a. Roland didn't think of himself as being superstitious
although he always carried his horseshoe (mascot, talisman).
He neither believed in horoscopes nor in fortune tellers.
- 30b. Roland didn't think of himself as being superstitious.
Although he always carried his horseshoe (mascot, talisman),
he neither believed in horoscopes nor in fortune tellers.

B Means and standard deviations for participants' overall recall accuracy in Experiments 1 to 4

Experiment 1 (auditory presentation, immediate recall): 89.36%, SD=5.4

Experiment 2 (RSVP, immediate recall): 82.91%, SD=7.6

Experiment 3 (RSVP, delayed recall): 71.64%, SD=10.97

Experiment 4 (RSVP, immediate recall): 88.54%, SD=6.9

C Results of Experiment 5 with all 30 items

C 1 Table of results

Table 9: Absolute frequencies and percentages (in brackets) of gender congruent and gender incongruent lure intrusions, correctly reproduced target words, and target omissions as a function of lure condition and state of processing in Experiment 5 (with all 30 items included)

	Target Word Part of the Penultimate Sentence			Target Word Part of the Ultimate Sentence		
	Without Lure	With Gender Congruent Lure	With Gender Incongruent Lure	Without Lure	With Gender Congruent Lure	With Gender Incongruent Lure
Gender Congruent Intrusions	19 (7.9%)	53 (22.1%)	14 (5.8%)	19 (7.9%)	47 (19.6%)	16 (6.7%)
Gender Incongruent Intrusions	9 (3.8%)	5 (2.1%)	35 (14.6%)	8 (3.3%)	3 (1.3%)	20 (8.3%)
Correct Target Reproductions	140 (58.3%)	130 (54.2%)	134 (55.8%)	171 (71.3%)	148 (61.7%)	157 (65.4%)
Target Omissions	72 (30.0%)	52 (21.7%)	57 (23.8%)	42 (17.5%)	42 (17.5%)	47 (19.6%)

C 2 Lure Intrusions

C 2.1 General analyses

C 2.1.1 Intrusion effect for gender congruent lure words

2-way interaction between lure condition and response type:

min. $LRCS1=17.30$, $df=1$, $p<.001$; min. $LRCS2=20.43$, $df=1$, $p<.001$

3-way interaction between, state of processing, lure condition, and response type:

all $LRCSs<1$, $df=1$, all $ps>.7$

3-way interaction between participants, lure condition, and response type: both $ps>.97$

3-way interaction between items, lure condition, and response type: both $ps>.14$

C 2.1.2 Intrusion effect for gender incongruent lure words

2-way interaction between lure condition and response type:

min. $LRCS1=8.87$, $df=1$, $p<.01$; min. $LRCS2=10.62$, $df=1$, $p<.01$

3-way interaction between state of processing, lure condition, and response type:

all $LRCSs<1.1$, $df=1$, all $ps>.3$

3-way interaction between participants, lure condition, and response type: both $ps>.99$

3-way interaction between items, lure condition, and response type: both $ps>.99$

C 2.1.3 Induced gender congruent intrusions vs. induced gender incongruent intrusions

2-way interaction between lure condition and response type:

min. $LRCS1=6.8$, $df=1$, $p<.01$; min. $LRCS2=8.5$, $df=1$, $p<.01$

3-way interaction between state of processing, lure condition, and response type:

both $LRCS1s<1$, $df=1$, both $ps>.4$; both $LRCS2s<1.5$, $df=1$, both $ps>.2$

3-way interaction between participants, lure condition, and response type: both $ps>.99$

3-way interaction between items, lure condition, and response type: both $ps>.12$

C 2.2 Separate analyses for the conditions in which the target word is part of the ultimate sentence

C 2.2.1 Intrusion effect for gender congruent lure words

2-way interaction between lure condition and response type:

min. $LRCS1=8.34$, $df=1$, $p<.01$; min. $LRCS2=9.75$, $df=1$, $p<.01$

3-way interaction between participants, lure condition, and response type: both $ps > .88$

3-way interaction between items, lure condition, and response type: both $ps > .88$

C 2.2.2 Intrusion effect for gender incongruent lure words

2-way interaction between lure condition and response type:

min. $LRCS1 = 2.42$, $df = 1$, $p > .12$; min. $LRCS2 = 2.81$, $df = 1$, $p > .09$

3-way interaction between participants, lure condition, and response type: both $ps > .99$

3-way interaction between items, lure condition, and response type: both $ps > .99$

C 2.2.3 Induced gender congruent intrusions vs. induced gender incongruent intrusions

2-way interaction between lure condition and response type:

min. $LRCS1 = 4.33$, $df = 1$, $p < .05$; min. $LRCS2 = 7.11$, $df = 1$, $p < .01$

3-way interaction between participants, lure condition, and response type: both $ps > .95$

3-way interaction between items, lure condition, and response type: both $ps > .95$

C 2.3 Separate analyses for the conditions in which the target word is part of the ultimate sentence

C 2.3.1 Intrusion effect for gender congruent lure words

2-way interaction between lure condition and response type:

min. $LRCS1 = 8.02$, $df = 1$, $p < .01$; min. $LRCS2 = 10.38$, $df = 1$, $p < .01$

3-way interaction between participants, lure condition, and response type: both $ps > .7$

3-way interaction between items, lure condition, and response type: both $ps > .7$

C 2.3.2 Intrusion effect for gender incongruent lure words

2-way interaction between lure condition and response type:

min. $LRCS1 = 6.59$, $df = 1$, $p < .05$; min. $LRCS2 = 8.52$, $df = 1$, $p < .01$

3-way interaction between participants, lure condition, and response type: both $ps > .95$

3-way interaction between items, lure condition, and response type: both $ps > .95$

C 2.3.3 Induced gender congruent intrusions vs. induced gender incongruent intrusions

2-way interaction between lure condition and response type:

min. $LRC S1=2.00$, $df=1$, $p<.17$; min. $LRC S2=2.33$, $df=1$, $p<.12$

3-way interaction between participants, lure condition, and response type: both $ps<.07$

3-way interaction between items, lure condition, and response type: both $ps<.1$

C 3 Recall accuracy

C 3.1 Proportion of correctly recalled words per clause

Table 10: Proportion of correctly recalled words per clause as a function of state of processing

	State of processing	
	[1,2] 3	1 [2,3]
Middle clause	71	78
Final clause	67	68

C 3.2 Proportion of clauses recalled perfectly

Table 11: Proportion of perfectly recalled clauses as a function of state of processing

	State of processing	
	[1,2] 3	1 [2,3]
Middle clause	48	51
Final clause	43	39

C 3.3 Clause recall: second vs. third clause

Correctly recalled words, middle clause part of the penultimate sentence:

$T=85187.5$, $z=-2.57$, $p<.05$

Correctly recalled words, middle clause part of the ultimate sentence:

$T=75240.5$, $z=-5.19$, $p<.001$

Perfectly recalled clauses, middle clause part of the penultimate sentence:

$T=21910$, $z=-1.81$, $p=.07$

Perfectly recalled clauses, middle clause part of the ultimate sentence:

$T=15288$, $z=-4.97$, $p<.001$

C 3.4 Clause recall: recall accuracy as a function of the state of processing

Correctly recalled words, middle clause: $T=71174.5$, $z=-3.96$, $p<.001$

Correctly recalled words, final clause: $T=96793$, $z=-.16$, $p>.8$

Perfectly recalled clauses, middle clause: $T=23152.5$, $z=-1.13$, $p>.25$

Perfectly recalled clauses, final clause: $T=21684$, $z=-1.87$, $p>.06$

D Instructions

D 1 Instructions for the experiments: German originals

D 1.1 Instruction for Experiment 1: German original

Liebe(r) Versuchsteilnehmer(in),

in diesem Experiment geht es um das Behalten von Sätzen. In jedem Durchgang des Experiments hörst Du zunächst eine Reihe von Wörtern. Um Deine Behaltensleistung zu erfassen, ertönt dann ein Satz. Diesen sprichst Du bitte nach, wenn es möglich ist wortwörtlich. Wichtig ist, dass Du den Satz so genau und so deutlich wie möglich wiedergibst. Nachdem Du den Satz wiedergegeben hast, drückst Du bitte die Leertaste des PC. Daraufhin hörst Du ein einzelnes Wort, von dem Du mit Tastendruck (SHIFT rechts JA, SHIFT links NEIN) sagen sollst, ob es in der zuvor gehörten Liste enthalten war oder nicht.

Damit ist der erste Durchgang beendet. Den nächsten Durchgang startest Du, indem Du die Leertaste drückst. Daraufhin werden Dir die nächste Liste und der folgende Satz vorgespielt. Den Satz sollst Du wieder nachsprechen. Dann hörst Du wieder das einzelne Wort. Dieser Ablauf wiederholt sich mehrfach bis zum Ende des Experiments. Bevor das Experiment beginnt, werden Dir sechs Probesätze dargeboten. Wenn Du noch Fragen hast, so stelle diese bitte jetzt. Ansonsten starte jetzt das Experiment, indem Du die Leertaste betätigst.

D 1.2 Instruction for Experiment 2 & 4: German original

Liebe(r) Versuchsteilnehmer(in),

in diesem Experiment geht es um das Behalten von Sätzen. In jedem Durchgang des Experiments siehst Du zunächst eine Reihe von Wörtern. Um Deine Behaltensleistung

zu erfassen, siehst Du dann einen Satz. Diesen sprichst Du bitte nach, wenn es möglich ist wortwörtlich. Wichtig ist, dass Du den Satz so genau und so deutlich wie möglich wiedergibst. Nachdem Du den Satz wiedergegeben hast, drückst Du bitte die Leertaste des PC. Daraufhin siehst Du ein einzelnes Wort, von dem Du mit Tastendruck (SHIFT rechts JA, SHIFT links NEIN) sagen sollst, ob es in der zuvor gehörten Liste enthalten war oder nicht.

Damit ist der erste Durchgang beendet. Den nächsten Durchgang startest Du, indem Du die Leertaste drückst. Daraufhin werden Dir die nächste Liste und der folgende Satz gezeigt. Den Satz sollst Du wieder nachsprechen. Dann siehst Du wieder das einzelne Wort. Dieser Ablauf wiederholt sich mehrfach bis zum Ende des Experiments. Bevor das Experiment beginnt, werden Dir sechs Probesätze dargeboten. Wenn Du noch Fragen hast, so stelle diese bitte jetzt. Ansonsten starte jetzt das Experiment, indem Du die Leertaste betätigst.

D 1.3 Instruction for Experiment 3: German original

Liebe(r) Versuchsteilnehmer(in),

in diesem Experiment geht es um das Behalten von Sätzen. In jedem Durchgang des Experiments siehst Du zunächst einen Satz und dann eine Reihe von Wörtern. Daraufhin siehst Du ein einzelnes Wort, von dem Du mit Tastendruck (SHIFT rechts JA, SHIFT links NEIN) sagen sollst, ob es in der zuvor präsentierten Liste enthalten war oder nicht. Danach sollst Du den Satz nachsprechen, wenn es möglich ist wortwörtlich. Wichtig ist, dass Du den Satz so genau und so deutlich wie möglich wiedergibst. Nachdem Du den Satz wiedergegeben hast, drückst Du bitte die Leertaste des PC. Damit ist der erste Durchgang beendet. Den nächsten Durchgang startest Du, indem Du die Leertaste drückst. Daraufhin werden Dir der folgende Satz und die nächste Liste präsentiert. Dann siehst Du wieder das einzelne Wort. Nach der Entscheidung per Tastendruck sollst Du wieder den Satz nachsprechen. Dieser Ablauf wiederholt sich mehrfach bis zum Ende des Experiments. Bevor das Experiment beginnt, werden Dir sechs Probesätze dargeboten.

Wenn Du noch Fragen hast, so stelle diese bitte jetzt. Ansonsten starte jetzt das Experiment, indem Du die Leertaste betätigst.

D 1.4 Instruction for Experiment 5: German original

Liebe(r) Versuchsteilnehmer(in),

in diesem Experiment hörst Du in jedem Durchgang einen kurzen Text, von dem Du Dir immer die jeweils letzten zwei Sätze merken sollst. Irgendwann wird der Text unterbrochen. Daraufhin hörst Du noch einmal den Anfang des jeweils vorletzten Satzes. Von diesem Punkt an sollst Du den Text dann bis zum Ende nachsprechen. Das heißt also, dass Du die beiden Sätze, die Du vor der Unterbrechung zuletzt gehört hast, wiedergeben sollst, und zwar möglichst wortwörtlich.

Dieser Ablauf wiederholt sich mehrfach. Bitte achte darauf, dass Du immer bei der Unterbrechung die letzten zwei Sätze möglichst genau und deutlich nachsprichst.

Um Dich mit dem Ablauf vertraut zu machen, werden Dir zunächst drei Probetexte dargeboten. Wenn Du nun keine Fragen mehr hast, kannst Du damit beginnen, indem Du die Leertaste drückst.

*D 2 English glosses**D 2.1 Instruction for Experiment 1: English glosses*

Dear participant,

This experiment is about sentence retention. In each trial, you are auditorily presented with a list of words and subsequently with a sentence. You have to orally repeat this sentence as verbatim as possible. Please press the spacebar after you have repeated the sentence. Then you will hear a single word. For this word, you have to decide via keystroke whether or not it was included in the word list (press the right shift key for yes and the left shift key for no).

This is the end of the first trial. You can start the next trial by pressing the space bar.

Then, you will be presented with the next word list and with the next sentence. Again, you have to repeat the sentence. And then you will again hear the single word.

This course is iterated several times. Before the experiment starts, you will be presented with six practice trials. You can start them now by pressing the space bar if you don't have any further questions.

D 2.2 Instruction for Experiment 2 & 4: English glosses

Dear participant,

This experiment is about sentence retention. In each trial, you will see a list of words and subsequently a sentence. You have to orally repeat this sentence as verbatim as possible. Please press the spacebar after you have repeated the sentence. Then you will see a single word. For this word, you have to decide via keystroke whether or not it was included in the word list (press the right shift key for yes and the left shift key for no).

This is the end of the first trial. You can start the next trial by pressing the space bar. Then, you will be presented with the next word list and with the next sentence. Again, you have to repeat the sentence. And then you will again see the single word.

This course is iterated several times. Before the experiment starts, you will be presented with six practice trials. You can start them now by pressing the space bar if you don't have any further questions.

D 2.3 Instruction for Experiment 3: English glosses

Dear participant,

This experiment is about sentence retention. In each trial, you will see a sentence and subsequently a list of words. Then you will see a single word. For this word, you have to decide via keystroke whether or not it was included in the word list (press the right shift key for yes and the left shift key for no). After this decision you have to orally repeat the sentence as verbatim as possible. Please press the spacebar after you have repeated the sentence.

This is the end of the first trial. You can start the next trial by pressing the space bar. Then, you will be presented with the next sentence and with the next word list. Again, you will see the single word. After your decision via keystroke you have to repeat the sentence.

This course is iterated several times. Before the experiment starts, you will be presented with six practice trials. You can start them now by pressing the space bar if you don't have any further questions.

D 2.4 Instruction for Experiment 5: English glosses

Dear participant,

In this experiment, you are in each trial auditorily presented with a short text of which you are to remember the last two sentences (each time). At any point, the text is interrupted. Then you once again hear the beginning of the penultimate sentence in each case. Starting from this point, you have to orally repeat the text up to its end. That is, you are expected to reproduce the two sentences you listened to lastly as verbatim as possible.

This course is iterated several times. Please make certain to always repeat the last two sentences as clearly and accurately as possible when the text is interrupted.

In order to become familiar with this process, you are first presented with three practice texts. You can start them now by pressing the space bar if you don't have any further questions.

*D 3 Instructions pre-tests: German originals**D 3.1 Instruction for the cloze test: German original*

Liebe Versuchsperson,

Vielen Dank für Deine Bereitschaft, an dieser Untersuchung teilzunehmen. Deine Aufgabe besteht darin, die folgenden Sätze so zu ergänzen, dass sie Dir sinnvoll erscheinen. Dabei ist es wichtig, dass Du nahe liegende Wörter einträgst, sie müssen nicht originell sein; schreib das auf, was Dir als erstes in den Sinn kommt. Für jedes Wort, das Du einfügen sollst, ist eine Linie gedruckt.

Beispiel:

Der Maler kleistert die Tapete großzügig ein, damit alles an der Wand haften bleibt.

D 3.2 Instruction for the semantic similarity rating: German original

In diesem Bogen sind jeweils zwei Wörter aufgeführt, die Du danach beurteilen sollst, wie ähnlich sie sich hinsichtlich ihrer Bedeutung sind. Bei großer Bedeutungsähnlichkeit der beiden Wörter vergibst Du eine hohe Punktzahl, wenn die Bedeutungen völlig unterschiedlich sind, eine niedrige.

Magazin	Zeitschrift	1 – 2 – 3 – 4 – 5 – 6 – 7 – 8 – 9 – 10
---------	-------------	--

Auto	Fahrzeug	1 – 2 – 3 – 4 – 5 – 6 – 7 – 8 – 9 – 10
------	----------	--

...

*D 4 Instructions pre-tests: English glosses**D 4.1 Instruction for the cloze test: English glosses*

Dear participant,

Thank you very much for your helpfulness. Your task is to complete the following sentences in a way that they make sense to you. It is important to choose words which suggest themselves; you do not need to be creative, just write down what first comes to mind. A line is printed for each word you are to fill in.

Example:

The painter pastes the wallpaper so that it sticks to the wall.

D 4.2 Instruction for the semantic similarity rating: English glosses

Dear participant,

In this rating, there are two words itemized in each case. You shall judge them for their mutual similarity with regard to their meanings. If they carry very similar meanings, you assign a high score and a low one if their meanings are completely different.

newspaper	journal	1 – 2 – 3 – 4 – 5 – 6 – 7 – 8 – 9 – 10
-----------	---------	--

car	vehicle	1 – 2 – 3 – 4 – 5 – 6 – 7 – 8 – 9 – 10
-----	---------	--

...

German Summary/Deutsche Zusammenfassung

I. Einleitung

Die vorliegende Arbeit befasst sich mit dem *verbalen Arbeitsgedächtnis*. In fünf Experimenten wird dabei am Beispiel von grammatischem Geschlecht untersucht, inwieweit das kurzfristige Behalten von sprachlicher Information auf denselben Repräsentationen beruht wie das Verstehen und Produzieren von sprachlichen Äußerungen. Zusammenfassend sollen an dieser Stelle einige Faktoren skizziert werden, die für das verbale Arbeitsgedächtnis als bedeutsam angesehen werden. Darüber hinaus wird das Modell, das dieser Arbeit zugrunde liegt, kurz beschrieben. Davon ausgehend werden die hier untersuchte Fragestellung und das dafür zentrale Paradigma näher erläutert. Es folgen ein Überblick über die Experimente und eine Zusammenfassung der Ergebnisse.

Das verbale Arbeitsgedächtnis umfasst die kognitiven Mechanismen, die genutzt werden, wenn gleichzeitig Informationen verarbeitet und sprachliche Informationen kurzfristig behalten werden müssen (siehe etwa Daneman & Carpenter, 1980). Die Frage, welche Mechanismen dem verbalen Arbeitsgedächtnis zugrunde liegen, wird kontrovers diskutiert. Ein wesentlicher Punkt, in dem sich verschiedene Auffassungen unterscheiden, ist der jeweils angenommene Grad der Trennung von Kurzzeitgedächtnis und Langzeitgedächtnis. Die aktuelle Entwicklung geht dahin, dass eine enge Beziehung zwischen kurzfristigem Behalten und den Repräsentationen des zu behaltenden Materials im Langzeitgedächtnis angenommen wird (vgl. etwa Cowan, 1995, 1999).

In der vorliegenden Arbeit sollen solche Modelle propagiert werden, in denen angenommen wird, dass das Verstehen von Sprache, das Produzieren von Sprache und das Behalten von Sprache die gleichen Repräsentationen nutzen (z.B. Dell, Schwartz, Martin, Saffran & Gagnon, 1997; Martin, Lesch & Bartha, 1999; Martin & Saffran, 1997). Eine solche Auffassung steht im Gegensatz zu solchen Modellen des Arbeitsgedächtnisses, die ein separates Arbeitsgedächtnissystem annehmen, das nur lose mit den jeweiligen Repräsentationen im Langzeitgedächtnis verbunden ist.

Als Evidenz für *sprachbasierte Modelle des Arbeitsgedächtnisses* lassen sich Befunde anführen, die zeigen, dass Faktoren, die auf den verschiedenen Ebenen sprachlicher Repräsentationen anzusiedeln sind, das Behalten beeinflussen. In diesem Zusammenhang wurden im Wesentlichen phonologische, lexikalische und se-

mentische Faktoren untersucht. Diese Faktoren lassen sich bereits auf Einzelwortebene manipulieren und zeigen dementsprechend einen Einfluss auf das serielle Behalten von Listen. Das serielle Listenbehalten ist die wohl am häufigsten verwendete Aufgabe zur Untersuchung des verbalen Arbeitsgedächtnisses (einen Überblick geben z.B. Baddeley, 1986, 1999; Eysenck & Keane, 2005).

Dabei zeigt sich etwa, dass Listen aus Wörtern, die ähnlich klingen, schlechter wiedergegeben werden als Listen aus Wörtern, die sich phonologisch unähnlich sind (*phonologischer Ähnlichkeitseffekt*, Conrad & Hull, 1964). Außerdem werden Listen aus Wörtern, die im Sprachgebrauch häufig sind (wie z.B. „Kuh“), besser wiedergegeben werden als Listen aus seltenen Wörtern (wie z.B. „Gnu“; Watkins, 1977). Dieser *Frequenzeffekt* wird auf den Einfluss des mentalen Lexikons (also eines Teil des sprachlichen Langzeitgedächtnisses) zurückgeführt. Ein semantischer Einflussfaktor ist die Unterscheidung konkreter Wörter (z.B. „Kuh“) und abstrakter Wörter (z.B. „Freiheit“). Hier zeigt sich eine bessere Wiedergabeleistung für konkrete Wörter, der so genannte *Konkretheitseffekt* (Walker & Hulme, 1999).

Nun ist aber offensichtlich, dass Sprache nicht nur aus einzelnen, unverbundenen Wörtern besteht. Die eigentliche Vielfalt sprachlicher Ausdrucksmöglichkeiten zeigt sich erst, wenn Wörter zu Sätzen und zu Texten verknüpft werden. So lassen sich aus den oben angeführten Wortbeispielen Sätze mit ganz unterschiedlicher Bedeutung erstellen (z.B. „Die Kuh beneidet das Gnu um seine Freiheit.“ vs. „Die Kuh betrügt das Gnu um ihrer Freiheit willen.“ vs. „Das Gnu schenkt der Kuh die Freiheit.“). Die unterschiedlichen Funktionen, die das Wort „Freiheit“ in diesen Sätzen erfüllt, lassen sich mittels unverbundener Wortlisten nicht vergleichen. Daraus ergibt sich die Frage, inwieweit sich Erkenntnisse, die aus der Untersuchung von Wortlisten gewonnen werden, auf komplexere kognitive Leistungen wie das Verstehen von Texten übertragen lassen.

Im Zentrum der vorliegenden Arbeit steht ein sprachbasiertes Modell des verbalen Behaltens, das besonders das kurzfristige Behalten von Sätzen adressiert (Rummer, 2003; Rummer & Engelkamp, in press; Rummer & Schweppe, in prep.). Essentiell für dieses Modell ist die Annahme einer wechselseitigen Abhängigkeit von Sprachverarbeitung und verbalem Behalten. Die Grundthese ist dabei, dass genau die Informationen kurzfristig gespeichert werden, die für die weitere Sprachverarbeitung (d.h., für das weitere Verstehen und/oder Produzieren von Sprache) potentiell *benötigt* werden, und zwar so lange, wie sie potentiell benötigt werden. Eine Repräsentation

tion der Bedeutung ist beispielsweise länger erforderlich als eine Repräsentation des genauen Wortlauts. Entsprechend dieser Annahme ergeben sich Behaltenseigenschaften direkt aus der Sprachverarbeitung.

Die vorliegende Studie untersucht die Modellannahmen anhand einer Informationsart, die bislang nicht als behaltenswirksam angesehen wurde, und zwar anhand von morphosyntaktischer Information, speziell anhand des grammatischen Geschlechts (oder auch *Genus*). Wie im Folgenden erläutert wird, sollte sich aufgrund der Rolle, die das Genus in der Sprachverarbeitung spielt, ein Einfluss von Genusinformation auch beim verbalen Behalten zeigen. Dies gilt zumindest dann, wenn man auf der Grundlage sprachbasierter Modelle des Arbeitsgedächtnisses argumentiert.

Im Deutschen ist das Genus bei vielen Wörtern relativ unabhängig vom natürlichen Geschlecht. Das folgende Beispiel illustriert diesen Umstand.

Als ich *das Auto*_[neut] kaufte, war *es*_[neut] erst ein Jahr alt. Dann habe ich *den Wagen*_[mask] fünf oder sechs Jahre gefahren und eigentlich nie Ärger mit *ihm*_[mask] gehabt. Voriges Jahr schließlich habe ich *die alte Kiste*_[fem] dann verkauft, weil *sie*_[fem] überall anfang zu rosten.

Darüber hinaus lässt sich an diesem Beispiel erkennen, dass Genusinformation Nomen und auf diese Bezug nehmende Pronomen miteinander verbindet. Beim Produzieren von Pronomen muss sicher gestellt sein, dass diese das gleiche grammatische Geschlecht haben, also *genuskongruent* sind. Beim Verstehen wird Genusinformation genutzt, um zu erkennen, auf welches vorher genannte Nomen sich ein Pronomen bezieht. Am obigen Beispiel wird auch deutlich, dass das Pronomen immer das Genus des zuletzt genannten Nomens annimmt, auch wenn sich mehrere Nomen mit unterschiedlichem grammatischem Geschlecht auf dieselbe Entität beziehen (es geht immer um dasselbe Fahrzeug).

Da Nomen und Pronomen oft durch mehrere Wörter, aber auch durch Teilsätze und Sätze getrennt sind, handelt es sich bei Genusinformation, ähnlich wie bei Bedeutungsinformation, um einen Informationstyp, der über einen längeren Zeitraum benötigt wird. Entsprechend der oben skizzierten Modellannahmen sollte Genusinformation also auch über einen längeren Zeitraum im Arbeitsgedächtnis verfügbar

sein und dementsprechend das Behalten beeinflussen. Dieser Frage soll in der vorliegenden Arbeit nachgegangen werden.

II. Fragestellung: Die Rolle von grammatischem Geschlecht beim Behalten von Sprache

Sprachbasierte Modelle des Arbeitsgedächtnisses machen die Vorhersage, dass Genusinformation beim Behalten von Sprache eine ähnliche Rolle spielt wie beim Verstehen und Produzieren von Sprache. In Sprachrezeption und –produktion zeigt sich ein solcher Einfluss dann, wenn (zumindest minimale) syntaktische Kontexte untersucht werden. Wie oben angedeutet, ist dies zum Beispiel bei der Interpretation anaphorischer Pronomen der Fall (z.B. Garnham, Oakhill, Ehrlich & Carreiras, 1995). Darüber hinaus werden lexikalische Fehler in der Sprachproduktion durch Genuskongruenz beeinflusst. Beispielsweise bleibt bei Versprechern häufig das grammatische Geschlecht erhalten (Marx, 1999), ebenso wie bei Benennfehlern von Personen mit aphasischer Sprachstörung (Kulke & Blanken, 2001). Für das verbale Behalten heißt das, dass Genusinformation dann einen Einfluss zeigen sollte, wenn syntaktisch strukturierte Informationen (also Sätze oder Texte) behalten werden müssen.

Um die Frage eines möglichen Behaltenseinflusses von Genusinformation zu untersuchen, kann ein bekanntes Paradigma aus der Forschung zum Satzbehalten herangezogen und modifiziert werden. Dieses Paradigma, das so genannte *Intrusionsparadigma* von Potter und Lombardi (1990), kombiniert die kurzfristige Wiedergabe von Sätzen mit dem Listenbehalten als Distraktoraufgabe. Dabei wird gezielt die lexikalische Selektion gestört. Jeder Satz enthält ein Zielwort, das inhaltlich nicht optimal passt. (z.B. „Der Chauffeur passte auf das *Fahrzeug* auf, während er auf seinen Chef wartete.“). Ein Wort, das stattdessen besser passt (im obigen Beispiel etwa „Auto“), ist in einer vor oder nach dem Satz präsentierten Distraktorliste entweder enthalten oder nicht. Bei der Wiedergabe des Satzes wird häufig das besser passende Distraktorwort fälschlicherweise eingesetzt („Der Chauffeur passte auf das *Auto* auf, während er auf seinen Chef wartete.“). Diese Intrusionen treten häufiger auf, wenn das Distraktorwort in der Liste präsentiert wurde. Potter und Lombardi (1990) führten das Intrusionsparadigma ein, um zu zeigen, dass Sätze nicht einfach aus dem Gedächtnis anhand einer Oberflächenrepräsentation abgerufen werden. Laut ihrer *Konzeptuellen Regenerierungshypothese* werden Sätze stattdessen aus der extrahierten

Satzbedeutung rekonstruiert, wobei vornehmlich solche Wörter verwendet werden, die kürzlich (also beim Satzverstehen) aktiviert wurden.

Im Sinne des Modells von Rummer (2003, s. auch Rummer & Engelkamp, in press) lässt sich die konzeptuelle Regenerierungshypothese jedoch erweitern. Demnach sollten neben der Satzbedeutung und lexikalischer Information auch andere Informationsarten behaltensrelevant sein, und zwar alle die Informationen, die beim Verstehen des Satzes extrahiert werden. Dazu gehören dann auch phonologische und morphosyntaktische Repräsentationen. Darüber hinaus sollten die einzelnen Informationsarten solange im Gedächtnis gehalten werden und damit solange behaltenswirksam sein, wie sie potentiell für die weitere Sprachverarbeitung genutzt werden können. Das Modell sagt damit vorher, dass Genusinformation bis zur Satz wiedergabe verfügbar sein sollte, da sie potentiell auch über die Satzgrenze hinaus benötigt wird.

Wie lässt sich nun das Intrusionsparadigma modifizieren, um einen Einfluss von Genuskongruenz beim Satzbehalten zu untersuchen? Die zu behaltenden Sätze lassen sich so konstruieren, dass statt *eines* Distraktorwortes gleich *zwei* solche Wörter besser in den Satz passen als das Zielwort. Eines dieser Distraktorwörter hat dann das gleiche grammatische Geschlecht wie das Zielwort (ist also genuskongruent), das andere hat ein anderes grammatisches Geschlecht (und ist somit genusinkongruent). Im oben genannten Beispiel („Der Chauffeur passte auf das Fahrzeug auf ...) ließen sich etwa Auto (genuskongruent) und Wagen (genusinkongruent) als Distraktorwörter einsetzen.

Wenn Genusinformation tatsächlich die lexikalische (Re-)Selektion beim Behalten beeinflusst, sollte die Präsentation des genuskongruenten Distraktorwortes häufiger zu Intrusionen führen als die des genusinkongruenten Distraktorwortes. In dieser Variante des Intrusionsparadigmas sind also beide Distraktoren dem Zielwort semantisch ähnlich, aber nur eins ist morphosyntaktisch ähnlich.

Dies ist das Hauptparadigma, das mit kritischen Modifikationen in den Experimenten 1 bis 4 eingesetzt wurde, um der Frage des Einflusses von Genusinformation auf das Behalten von Sätzen nachzugehen. Experiment 5 erweitert den Skopus auf das Behalten von kurzen Texten (Näheres unten).

III. Experimente

Experiment 1. In Experiment 1 wurde die oben skizzierte Variante des Intrusionsparadigmas eingesetzt. Dabei wurde den Versuchspersonen zunächst über Lautsprecher

eine Distraktorliste dargeboten, die entweder das genuskongruente Distraktorwort, das genusinkongruente Distraktorwort oder ein unrelatiertes Kontrollwort enthielt. Im Anschluss daran wurde der Satz präsentiert, ebenfalls auditiv. Unmittelbar im Anschluss an die Darbietung des Satzes sollte dieser möglichst wörtlich wiedergegeben werden. Schließlich wurde ein einzelnes Wort präsentiert. Die Versuchspersonen sollten entscheiden, ob dies Teil der eingangs präsentierten Wortliste war.

Das Material bestand aus Sätzen, die analog zu dem oben angeführten Beispiel konstruiert waren. Die beiden Typen von Distraktorwörtern (genuskongruente und genusinkongruente) wurden vorab unter anderem hinsichtlich ihrer Adäquatheit im Satz, ihrer semantischen Ähnlichkeit zum Zielwort und ihrer Wortlänge kontrolliert.

Es zeigte sich, dass die Präsentation des kongruenten Distraktorwortes zu mehr entsprechenden Intrusionen führte als die Präsentation des inkongruenten Distraktorwortes. Dieses Ergebnis spricht für die Annahme, dass Genusinformation das Satzbehalten beeinflussen kann.

Experiment 2. Allerdings ist nicht auszuschließen, dass der hier beobachtete Effekt (im Folgenden *Genuskongruenzeffekt*) möglicherweise nicht auf morphosyntaktischer Information beruht, sondern auf phonologischer Information. Die Genusinkongruenz schlägt sich nämlich auch in der Form des jeweiligen Artikels nieder. Wird „das Fahrzeug“ durch „das Auto“ ersetzt, bleibt der Artikel gleich; wird stattdessen „der Wagen“ eingesetzt, besteht eine zusätzliche Diskrepanz zwischen den Artikeln. In Experiment 2 wurde deshalb die Bereitstellung phonologischer Information bei der Darbietung reduziert. Statt über Lautsprecher wurden nun Sätze und Listen mittels *rapid serial visual presentation* (RSVP) dargeboten. Die Präsentation erfolgte dabei Wort für Wort in der Mitte des Bildschirms bei einer Präsentationszeit von 200 ms pro Wort. Es wurde angenommen, dass diese Präsentationszeit zu schnell für eine adäquate phonologische Enkodierung der Wörter ist (vgl. Potter, 1999). In allen übrigen Punkten entsprach das Vorgehen dem des ersten Experiments.

Auch die Ergebnisse entsprachen denen aus Experiment 1. Selbst bei Darbietung mittels RSVP induzierte die Präsentation des genuskongruenten Distraktorwortes mehr genuskongruente Intrusionen als die Präsentation des genusinkongruenten Distraktorwortes genusinkongruente Intrusionen.

Experiment 3. Allerdings ist die Befundlage zur Verfügbarkeit phonologischer Information bei RSVP nicht eindeutig. Einige Autoren vertreten die Ansicht, dass selbst unter diesen Darbietungsbedingungen noch phonologische Information bereitgestellt wird (z.B. Coltheart, 1999a, 1999b). Eine Möglichkeit, die Verfügbarkeit phonologischer Information weiter zu reduzieren, bietet die Verlängerung des Zeitraums zwischen Satzdarbietung und –wiedergabe, da phonologische Information rasch zerfällt. Durch eine Veränderung der Abfolge lässt sich eine „*delayed recall*“-Situation herstellen (s. auch Rummer & Engelkamp, 2003). In diesem Fall wird zunächst der Satz präsentiert und dann die Liste. Darauf folgt die Rekognitionsentscheidung bezüglich des präsentierten Einzelwortes und erst dann muss der Satz wiedergegeben werden. Eine solche Vorgehensweise reduziert die Verfügbarkeit phonologischer Information in zweierlei Hinsicht, zum einen durch die Darbietung mittels RSVP, zum anderen durch den verzögerten Abruf beim *delayed recall*. Diese Bedingungen wurden in Experiment 3 umgesetzt, das übrige Vorgehen entsprach dem der ersten beiden Experimente.

Insgesamt ergaben sich in diesem Experiment mehr Intrusionen als bei unmittelbarer Wiedergabe, der vorhergesagte Unterschied zwischen genuskongruenten und genusinkongruenten Intrusionen blieb aber erhalten.

Experiment 4. Es besteht jedoch weiterhin die Möglichkeit, dass eine *lexikalische* Repräsentation des Zielwortartikels für den Genuskongruenzeffekt verantwortlich ist. Wie oben erwähnt, besteht bei „das Fahrzeug“ und „der Wagen“ zusätzlich zur Genusinkongruenz eine Diskrepanz zwischen den Artikeln. Behielte man eine lexikalische Repräsentation des Artikels „das“ im Gedächtnis, würde allein das die Wahrscheinlichkeit des Auftretens von Intrusionen durch „Wagen“ verringern. Das Deutsche bietet allerdings die Möglichkeit, eine *Genusinkongruenz* bei gleichzeitiger *Artikelkongruenz* zu realisieren. Macht man aus „Der Chauffeur passt auf das Fahrzeug auf“ „Der Chauffeur wartet neben dem Fahrzeug“, bliebe auch beim Ersetzen von „Fahrzeug“ durch „Wagen“ der Artikel gleich („Der Chauffeur wartet neben dem Wagen“ im Gegensatz zu „Der Chauffeur passt auf den Wagen auf“). Es lässt sich also zusätzlich die Eindeutigkeit der Genusmarkierung am Artikel manipulieren. Eine solche Manipulation wurde in Experiment 4 vorgenommen. Das übrige Vorgehen entsprach dem in Experiment 2.

Wiederum zeigte sich ein Genuskongruenzeffekt: Es traten also mehr induzierte genuskongruente Intrusionen als induzierte genusinkongruente Intrusionen auf. Dieses Muster wurde nicht beeinflusst durch die Eindeutigkeit der Genusmarkierung. Numerisch war der Unterschied zwischen genuskongruenten und genusinkongruenten Intrusionen bei ambiger Genusmarkierung sogar größer. Die Ergebnisse sprechen dafür, dass tatsächlich Genusinformation den Genuskongruenzeffekt verursacht.

Damit sprechen die Ergebnisse der Experimente 1 bis 4 dafür, dass das grammatische Geschlecht beim Behalten von Sätzen eine vergleichbare Rolle wie bei der Sprachverarbeitung spielt.

Experiment 5. Obgleich gezeigt wurde, dass Genusinformation beim Behalten von Sätzen genutzt wird, bleibt weiterhin die Frage nach der *Dauer* bzw. Beständigkeit des Einflusses von Genusinformation offen. Experiment 3 weist auf einen Einfluss über die unmittelbare Satz wiedergabe hin. Zusätzlich zum zeitlichen Abstand lässt sich aber auch die „syntaktische Distanz“ untersuchen.

Bereits in den späten 1960er und den 1970er Jahren wurden Untersuchungen durchgeführt, die sich mit der Frage beschäftigten, welche sprachlichen Repräsentationen für das Behalten von Sätzen relevant sind und für wie lange unterschiedliche Arten von Repräsentationen behalten werden. Im Folgenden werde ich kurz auf die in diesem Zusammenhang von Jarvella (1971, 1979) durchgeführten Untersuchungen zum kurzfristigen Behalten von Texten näher eingehen, da das entsprechende Paradigma in modifizierter Form im abschließenden Experiment der vorliegenden Arbeit verwendet wurde.

Jarvella (1971) präsentierte seinen Versuchspersonen Texte, die während der Darbietung immer wieder unterbrochen wurden. Im Anschluss an diese Unterbrechungen sollten die Versuchspersonen so viel wie möglich vom zuvor gehörten Text wiedergeben. In den kritischen Durchgängen bestanden die jeweils letzten zwei Sätze vor der Unterbrechung aus drei Teilsätzen, die zu unterschiedlichen syntaktischen Strukturen kombiniert werden konnten. Entweder bildeten die letzten beiden Teilsätze einen gemeinsamen Satz oder die ersten beiden Teilsätze. Der mittlere Teilsatz gehörte dementsprechend entweder zum zuletzt dargebotenen Satz oder zum vorangegangenen.

Jarvella nahm an, dass die Oberfläche des letzten Satzes zum Zeitpunkt der Wiedergabe besser verfügbar sein sollte als die Oberfläche des vorangegangenen

Satzes und fand diese Annahme bestätigt. Der mittlere Teilsatz wurde in Abhängigkeit von seinem Verarbeitungsstatus (Teil des letzten vs. Teil des vorletzten Satzes) unterschiedlich genau wiedergegeben, trotz gleichen Wortlauts. Gehörte der mittlere Teilsatz zum vorletzten Satz des Textes, wurde er wesentlich schlechter wiedergegeben als der letzte Teilsatz. Gehörte er aber zum letzten Satz, näherte sich die Wiedergabeleistung der des letzten Teilsatzes an. Daraus lässt sich schließen, dass Oberflächeninformation im Wesentlichen für den zuletzt verarbeiteten Satz verfügbar ist. Jarvella nimmt an, dass an Satzgrenzen solche Information zugunsten einer reicheren, konzeptuellen Enkodierung verloren geht.

Da das grammatische Geschlecht auch als Oberflächeninformation anzusehen ist, sollte diese entsprechend der Annahmen von Jarvella ebenfalls an Satzgrenzen verloren gehen. Das hier zugrunde gelegte Modell sagt dagegen vorher, dass sich ein Einfluss von Genusinformation auch über Satzgrenzen hinaus zeigen sollte. Dies ergibt sich daraus, dass für die Interpretation von Pronomen mit anaphorischer Referenz eine Repräsentation des grammatischen Geschlechts auch über Satzgrenzen hinaus *benötigt* wird.

In Experiment 5 wurde der Frage nachgegangen, ob Genusinformation tatsächlich über die Satzgrenzen hinaus das verbale Behalten beeinflusst. Dieses letzte Experiment der vorliegenden Arbeit erweitert den Skopus somit auf das Textbehalten. Dazu wurde Jarvellas Paradigma zum Textbehalten mit der Genusvariante des Intrusionsparadigmas kombiniert. Auf diese Weise lässt sich das Auftreten des Genuskongruenzeffekts in Abhängigkeit vom syntaktischen Status des Teilsatzes, in dem sich das Zielwort befindet, beziehungsweise in Abhängigkeit von der „syntaktischen Distanz“ zwischen der Präsentation des Zielworts und der Wiedergabe untersuchen.

Das Material war dabei folgendermaßen aufgebaut: Einzelne kurze Texte bestanden je aus einer Kontextpassage und einer Zielregion. Die Zielregion bestand wie bei Jarvella aus drei Teilsätzen, die auf zwei Arten kombiniert werden konnten: Entweder bildeten Teilsatz 1 und Teilsatz 2 einen gemeinsamen Satz oder Teilsatz 2 und Teilsatz 3 gehörten zusammen. Im Unterschied zu Jarvellas Material enthielt in der hier eingeführten Version der kritische zweite Teilsatz ein Zielwort analog zu den Einzelsatzexperimenten. Anders als bei Sätzen wurde hier das genuskongruente oder genusinkongruente Distraktorwort (bzw. das unrelatierte Kontrollwort) nicht in einer separaten Liste, sondern in dem Satz präsentiert, der der Zielregion direkt voranging.

Auf diese Weise konnten sowohl der syntaktische Status des kritischen Teilsatzes als auch das Distraktor- bzw. Kontrollwort manipuliert werden. Das folgende Beispiel veranschaulicht das verwendete Material sowie die Manipulationen.

Kontextpassage

Denise hatte sich für den Samstag fest vorgenommen, einen Stadtbummel zu machen. Am frühen Morgen schon ließ sie sich mit ihrem Auto_[genuskongruentes Distraktorwort]/Wagen_[genusinkongruentes Distraktorwort]/Hund_[Kontrollwort] abholen und nahe der Fußgängerzone absetzen.

Zielregion

- (a) Der Parkplatz füllte sich langsam,
während der Chauffeur das Fahrzeug_[Zielwort] bewachte.
Denise ging zum Einkaufen in die Innenstadt.
- (b) Der Parkplatz füllte sich langsam.
Während der Chauffeur das Fahrzeug_[Zielwort] bewachte,
ging Denise zum Einkaufen in die Innenstadt.

Wie oben erwähnt, sollte sich in beiden syntaktischen Varianten („Teilsatz 1 + Teilsatz 2. Teilsatz 3.“ und „Teilsatz 1. Teilsatz 2 + Teilsatz 3.“) ein Genuskongruenzeffekt zeigen. Die Ergebnisse entsprechen dieser Annahme weitgehend, aber nicht ganz. Insgesamt war wie schon in den Experimenten zum Satzbehalten ein Genuskongruenzeffekt zu beobachten. Darüber hinaus ergaben separate Analysen für beide Syntaxbedingungen einen Genuskongruenzeffekt, wenn der kritische Teilsatz und damit das Zielwort zum letzten Satz gehörte und einen marginal signifikanten Genuskongruenzeffekt, wenn der kritische Teilsatz und damit das Zielwort zum vorletzten Satz gehörte. Die Ergebnisse aus Experiment 5 sprechen für einen Einfluss auch über die Satzgrenzen hinweg.

IV. Zusammenfassung der Ergebnisse und abschließende Bemerkungen

Insgesamt zeigen die Ergebnisse, dass Genusinformation tatsächlich das kurzfristige Behalten von syntaktisch strukturiertem sprachlichem Material beeinflusst. Bisher wurden Behaltenseinflüsse nur für phonologische und lexikalisch-semantische Faktoren gezeigt. Hier ergab sich also zum ersten Mal ein Einfluss für morphosyntaktische Information. Ein solcher Einfluss entspricht den Annahmen sprachbasierter Modelle

des Arbeitsgedächtnisses. Außerdem fanden sich Hinweise darauf, dass der Einfluss von Genusinformation über Satzgrenzen hinaus wirksam ist. Dies entspricht der darüber hinausgehenden Annahme, dass die Behaltensdauer von den Erfordernissen der Sprachverarbeitung abhängt. In Bezug auf diesen Punkt stellen die vorliegenden Befunde einen ersten Hinweis dar. Eine solche Annahme muss aber noch für andere Arten von Repräsentationen untersucht werden. Die Rolle von Genusinformationen beim Behalten von Sätzen und kurzen Texten weist also darauf hin, dass die Verarbeitung und das Behalten von Sprache tatsächlich enger verbunden sind als es viele Modelle annehmen. Solche Befunde passen auch zur zunehmenden Fokussierung auf die Rolle von im Langzeitgedächtnis gespeicherter Informationen bei Arbeitsgedächtnisaufgaben (s. z.B. Baddeley, ab 2000; Cowan, 1995, 1999).