The moment in between:
Planning speech while listening

Svetlana Gerakaki

The educational component of the doctoral training was provided by the International Max Planck Research School (IMPRS) for Language Sciences. The graduate school is a joint initiative between the Max Planck Institute for Psycholinguistics and two partner institutes at Radboud University – the Centre for Language Studies, and the Donders Institute for Brain, Cognition and Behaviour. The IMPRS curriculum, which is funded by the Max Planck Society for the Advancement of Science, ensures that each member receives interdisciplinary training in the language sciences and develops a well-rounded skill set in preparation for fulfilling careers in academia and beyond.

More information can be found at www.mpi.nl/imprs.

ISBN: 978-94-92910-13-4

Cover design: Panos Gerakakis

Printed and bound by Ipskamp Printing, Nijmegen

The research reported in this thesis was conducted at the Max Planck Institute for Psycholinguistics, in Nijmegen, the Netherlands.

The moment in between:
Planning speech while listening

Proefschrift
ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op woensdag 8 juli 2020
om 12.30 uur precies

door
Svetlana-Lito Gerakaki
geboren op 1 april 1979
te Athene (Griekenland)

Promotor:
Prof. dr. P. Hagoort

Copromotor:
Dr. M.J. Sjerps (Ministerie van Onderwijs, Cultuur en Wetenschap)

Manuscriptcommissie:
Prof. dr. J. M. McQueen
Dr. J. A. M. Holler
Dr. R. M. Willems

The moment in between:
Planning speech while listening

Doctoral Thesis

to obtain the degree of doctor
from Radboud University Nijmegen
on the authority of the rector magnificus prof. dr. J.H.J.M. van Krieken,
according to the decision of the Council of Deans
to be defended in public on Wednesday, July 8, 2020
at 12.30 hours

by
Svetlana-Lito Gerakaki
born on April 1 1979
In Athens (Greece)

Supervisor:
Prof. dr. P. Hagoort

Co-supervisor:
Dr. M.J. Sjerps (Ministry of Education, Culture and Science, the Netherlands)

Doctoral Thesis Committee:
Prof. dr. J. M. McQueen
Dr. J. A. M. Holler
Dr. R. M. Willems

Table of Contents

Chapter 1

Introduction

Until very recently, studies in psycholinguistics typically adhered to a strict distinction between research on language comprehension and research on language production. These lines of research have led to a dramatic increase in our understanding of the human language system (e.g. Costa & Santesteban, 2004; Ferreira & Bailey, 2004; Hagoort, Hald, Bastiaansen, & Petersson, 2004; Indefrey & Levelt, 2004; Just, Carpenter, & Keller, 1996; Levelt, 1999, 2001; Levelt, Roelofs, & Meyer, 1999; Meyer, Ouellet, & Häcker, 2008; Smith & Wheeldon, 1999). However, when I started the research for this thesis in 2011 (all research reported in this thesis was conducted between 2011 and 2014) the field of psycholinguistics was becoming increasingly aware that a true understanding of language behavior also requires studying language in its natural environment. Hence, the field had started to shift towards studies investigating how comprehension and production work together as a system (e.g. Kempen, Olsthoorn, & Sprenger, 2011; Kubose, Bock, Dell, Garnsey, Kramer, & Mayhugh, 2006; Menenti, Gierhan, Segaert, & Hagoort, 2011; Menenti, Pickering, & Garrod, 2012; Roelofs, Ozdemir, & Levelt, 2007), as well as towards studies on conversational interaction (see for example Levinson, 2015; and the frontiers research topics special issue on turn-taking in human communicative interaction edited by Holler, Kendrick, Casillas, & Levinson, 2016). In an effort to contribute to this relatively new direction in language research, the focus of this thesis is on the interface between listening and concurrent planning and the processing constraints that might arise from it. A better understanding of the interplay between speech production and perception is critical for gaining better insight into the processing constraints that shape language behavior in everyday conversational interaction.

Turn-taking timing in conversation

A conversational interaction includes at minimum two interlocutors that take turns as the conversation unfolds. The act of switching between the roles of listener and speaker has been labeled turn-taking. Turn-taking timing, then, refers to the duration of the interval between the end of a speaker's utterance (turn-ending) and the initiation of the other speaker's utterance. Research on turn-taking timing suggests that interlocutors are very good at coordinating turns. For example, Ten Bosch, Oostdijk, and Boves (2005) reported a median gap between turns of only 330 ms (see also Beattie & Barnard, 1979; de Ruiter, Mitterer, & Enfield, 2006; Sacks, Schegloff, & Jefferson, 1974; Scott, Mcgettigan, & Eisner, 2009; Stivers et al., 2009; Wilson & Wilson, 2005).

These rapid turns give rise to the subjective experience that turn-taking timing is typically very smooth and easy (e.g. Garrod & Pickering, 2004; Pickering & Garrod, 2004; Scott et al., 2009; Stivers et al., 2009). At the same time language production research has shown that even single picture-naming can take 600 ms (Indefrey, 2011; Indefrey & Levelt, 2004), and preparing the first phrase of a sentence can easily take up more than a second (Konopka, 2011). So given the reported gap durations on turn-taking timing, how can interlocutors prepare and produce such a timely response? This phenomenon has been labeled *"the core psycholinguistic puzzle"* (Levinson & Torreira, 2015). How is this smooth timing accomplished, despite the time needed to prepare speech?

CHAPTER 1: INTRODUCTION

Two main processes have been proposed to contribute to smooth turn-taking-timing; 1) anticipating cues signaling turn-endings and 2) early speech planning onset. The idea that taking over a turn timely involves projecting when the incoming turn will end, and initiating production processes before the incoming turn is completed, were already part of Sacks and colleagues theory on conversational turn-taking (Sacks et al., 1974). That is, early on in conversational research it was acknowledged that the characteristics of conversational interaction have consequences for language processing.

Yet again the reverse is also true, as conceiving a psychologically real turn-taking model requires taking into account the constraints imposed by psycholinguistic processing (Levinson and Torreira, 2015). The Levinson-Torreira Model (Levinson & Torreira, 2015) embraces Sacks and colleagues (1974) theory, while using current findings from psycholinguistic research and statistical studies of corpora to make further assumptions on the cognitive processes involved in successful turn-taking. According to this model comprehenders use prediction to extract the speech act of the incoming utterance as soon as possible. Once this is done speech planning for the reply can initiate, while still monitoring the incoming speech signal for turn-final cues. Depending on when in time the turn-final cues are perceived, the already prepared response might have to be buffered, until time for articulating has arrived. As a result fast turn-taking is accomplished.

In this thesis I focus on three aspects of the Levinson-Torreira turn-taking Model. First, I investigate the idea of early speech planning (that is while still listening; chapter 2). In doing so, I further investigate whether planning efficiency can also contribute to fast turn-taking-timing. Second, I go over to examine how overlap in listening and speech planning might affect listening quality (chapters 3, 4 and 5).

The idea of early speech planning implies that maintaining the fast pace of turn-taking requires that comprehension and production must at least at times happen at the same time. That is, a form of dual-tasking is part of a successful turn-taking-timing. But how well can interlocutors perform in this dual tasking situation? Are there processing costs arising for speech perception and possibly even speech production? Finally, I consider the role of prediction in coping with the overlapping processes of speech perception and production (chapters 4 and 5). That is, rather than focusing on the role of prediction in extracting the speech act of the incoming utterance and possibly even providing information for turn-end completion (as proposed by the Levinson-Torreira model), I investigate how prediction might affect performance, mainly in speech perception, but also in speech planning, when these processes have to happen in overlap. In the following sections I further elaborate on these aspects and introduce a number of questions related to them.

How do interlocutors anticipate turn-endings? One of the two main processes that have been proposed to contribute to smooth turn-taking-timing is turn-end anticipation. That is, a listener can use information received from the speaker, to anticipate when that turn is going to end. This way the next speaker can start talking at the projected turn ending. The *projection theory* posits that anticipation or projection of a turn end is based on structural and contextual information (Sacks et al., 1974). That is, turn units that are syntactic in nature in combination with prosodic cues, function as Transition Relevant Places (TRP).

Cues to indicate upcoming turn-ends can be derived from interlocutors facial expressions, like information from eye-gaze (e.g. Duncan & Fiske, 1977; Kendon, 1967), from pragmatic information (Ford & Thompson, 1996) or directly from the speech signal. For example, it has been suggested that a high pitch peak can signal that the turn will end at the next syntactic completion point, thereby allowing for ample time to anticipate a turn's ending (e.g., Schegloff 1996; more on prosody also by Couper-Kuhlen & Selting, 1996; for cues related to syntactic structure and prosody see de Ruiter and colleagues, 2006).

Most importantly, the cues that have been implicated seem to work in additive ways and some have proven more useful than others. For example, Schaffer (1983) concluded that lexicosyntactic information, compared to intonation, offers a more reliable source to anticipate a turn's ending. A similar conclusion was drawn by de Ruiter and colleagues (2006) who looked into how lexico-syntactic content and intonation affect turn-end prediction in recordings of natural conversations. In a follow up study Magyari and de Ruiter (2008) provided some evidence that the advantage of lexico-syntactic content is that interlocutors can project a turn-ending because they can predict the content of the turn-final words. That is, by predicting the turn-final words, interlocutors also know when that turn will end (see also Magyari, Bastiaansen, de Ruiter, & Levinson, 2014; Magyari, de Ruiter, & Levinson, 2017).

The role of turn-end anticipation is not a central topic of this thesis. Even so, it is important to introduce the main ideas behind it, since a lot of evidence seems to suggest that anticipating or projecting a turn's end is a central aspect of fast turn-taking and an essential part of any cognitively valid turn-taking model (for more research on the topic see Bögels & Torreira, 2015; Levinson & Torreira, 2015; Riest, Jorschick, & de Ruiter, 2015). Chapters 4 and 5 do provide some links to how turn-end anticipation might affect turn-taking-timing. These will be discussed in more detail in chapter 6.

But are there other factors that affect turn-taking timing? Recently, it was proposed that the facilitation in turn-taking due to content prediction, does not result from allowing a successful turn-end projection (see Magyari et al., 2014, 2017; Magyari & de Ruiter, 2008), but rather from assisting in preparing a timely response (Corps, Crossley, Gambi, & Pickering, 2018). To be able to actually take over the turn with rapid timing, one also needs to rapidly start preparing a reply. That is, only when a response has been at least partly planned in advance will an interlocutor have something ready to articulate once the turn-final cues are perceived. In line with Corps and colleagues (2018), Barthel, Meyer, and Levinson (2017) concluded that response turn preparation and the timing of its articulation should be seen as separate processes in turn-taking models, because picking up on turn-final cues assists interlocutors in initiating their response faster, while it does not affect the timing of speech planning initiation (Barthel et al., 2017). For response turn preparation to begin, the incoming turn's message has to be sufficiently recognized. Thus, there are at least two sources affecting turn-taking timing; planning initiation- and response initiation- timing, with the former linking to advance planning processes and the later to turn-end anticipation processes. Even so, it is reasonable to assume that these separate processes are related in as far as response initiation cannot begin if response turn preparation has not reached a sufficient level that would allow for a reply chunk to be articulated. Therefore, fast turn-taking-timing seems to involve both turn-end anticipation and advance response planning.

Do interlocutors plan speech while listening to speech? If the answer is yes, interlocutors would have to start planning while still listening to the other speaker, in order to maintain a fast and smooth turn-taking. This is one of the central ideas proposed in the turn-taking model of Levinson and Torreira (2015). They suggest that conversational partners can often respond with short latencies because they are preparing their response as soon as possible (i.e., while still listening to the incoming turn). That is, it is suggested that the production system may begin to formulate a response even at the level of the phonological form (see also Bögels, Magyari, & Levinson, 2015), but with the actual articulation held in a buffer, waiting for execution until the comprehension system signals completion (or near-completion) of the incoming turn.

In chapter 2 of this thesis I investigated whether indeed fast turn-taking involves initiating speech planning while still listening to the incoming turn. To this end, I assessed how the onset- timing of speech planning in a scene description task is modulated by speech input that allows for early as opposed to late speech planning initiation. While the findings reported in Chapter 2 shed some light on this question, it is important to note that after these experiments were carried out, other research has also shown that, at the very least, interlocutors start planning their turn shortly before the turn-end of their conversational partner (Bögels, Casillas, & Levinson, 2018; Bögels, Magyari, & Levinson, 2015; Boiteau, Malone, Peters, & Almor, 2014; Sjerps & Meyer, 2015). Yet, at the time the experiments for Chapter 2 were conducted (2011-2012), no direct evidence existed on whether interlocutors do indeed plan while listening. Thus it was important to first establish whether this was indeed the case (chapter 2), in order to go on and study the interface between listening and concurrent planning in the subsequent reseach projects (described in Chapters 3, 4, 5).

Can differing demands in response planning independently contribute to turn-taking-timing? Even when interlocutors start planning their reply early, differing demands on the planning process might render it more or less successful. That is, the planning of one's response is not always equally demanding. As a result, preparing a response can be more or less successful and affects when in time a first portion of the reply will be available to articulate. I use the term "efficiency of planning" to describe this situation. For example, having difficulties deciding what message one wishes to convey can delay a response (see for example Pope, Blass, Cheek, & Bradford, 1971; Siegman, & Pope, 1972). But there are also lower-level aspects of speech production that could affect planning and speaking onset, such as the length of the response to be produced. A single-word reply should usually take less time to produce than uttering a whole sentence. Similarly, Smith and Wheeldon (1999) found that planning initial complex noun phrases results in significantly longer onset latencies than planning initial simple noun phrases. Another potential factor determining whether a timely response will be produced might link to conceptualization difficulties. The contribution of efficiency of planning to turn-taking timing is the second question addressed in chapter 2.

If fast turn-taking-timing requires dual tasking, are there trade offs between listening and speech planning? Planning onset timing and planning efficiency describe possible factors entering the turn-taking timing equation and refer to production processes. Given the Levinson & Torreira model, an interlocutor does not prepare her response in isolation, nor does she merely listen to anticipate turn-ends. One of the main theoretical puzzles of language use in conversation is the fact that turn taking is usually fast, while speech planning usually takes much longer. Solving this puzzle seems to involve some form of dual-tasking. That is, while an interlocutor listens, in order to process the incoming speech signal, she is also planning her response. It thus seems that the interlocutor is a multi-tasker that has to successfully share cognitive resources between the tasks of listening (anticipating included) and preparing her response. This may be a highly complex dual-task scenario especially since speech comprehension and production rely on much of the same neural circuitry (e.g., Segaert, Menenti, Weber, Petersson, & Hagoort, 2012).

Such a need to share domain-general cognitive resources (e.g. Baddeley, 1976; Becic et al., 2010; Kemper, Herman, & Lian, 2003; Kemper, Schmalzried, Herman, & Mohankumar, 2011; Lavie, Hirst, Fockert, & Viding, 2004; Lavie, 2005; Meyer & Kieras, 1997; Pashler, 1984, 1994) could trigger an interaction of speech perception and production processes. In particular, an extensive dual tasking literature posits that when a limited amount of cognitive capacity is distributed across multiple tasks, performance in the two tasks typically involves a direct tradeoff. That is, increases in performance on one task are correlated with decreases in the secondary task (Somberg & Salthouse, 1982). In research looking into how secondary nonlinguistic tasks might affect concurrent speech perception, it has been demonstrated that such interference is pervasive across different levels of linguistic processing. For example, this relation has been evident in effects on the perception of individual phonemes (e.g. Gordon, Eberhardt, & Rueckl, 1993; Mattys & Wiget, 2011), individual words (e.g. Cleland, Tamminen, Quinlan, & Gaskell, 2012), and the perception of sentences (Bosker, Reinisch, & Sjerps, 2017).

Importantly, research has provided evidence of such interference effects in fairly naturalistic linguistic settings. For example, it has been shown that driving a car reduces concurrent language production and comprehension performance as measured by retelling performance, story comprehension and long-term memory encoding of heard stories (Becic et al., 2010). And nonlinguistic task performance is similarly affected by language behavior: language production and perception tasks have a strong impact on several aspects of concurrent driving (e.g. Kubose et al., 2006; Strayer & Johnston, 2001).

So how do interlocutors cope with this multi-task setting in turn-taking? Is there a tradeoff/interference between listening and planning to speak? Even though for the multi-tasking interlocutor the goal of successful participation in conversation would not allow for any trade-off effects to occur, it is quite likely that they do occur. And when interference does occur, the costs in processing can arise in both listening as well as speech production. The cost arising on the listening side is the main topic of chapters 3, 4 and 5, in which I investigate whether listening performance drops, due to the need to prepare speech. The cost arising on the production side is partly addressed in chapter 3, but is not a central topic of this thesis.

Can predictability help listeners cope with those dual-task demands? The potentially interfering dual-task effects in turn-taking raise the question what mechanisms interlocutors may have at their disposal to help cope with these pressures. Could prediction help interlocutors cope with the dual-task demands? In the relevant literature prediction has mostly been discussed as a tool in assisting comprehension, thereby supporting a timely planning initiation; and as a tool in assisting turn-end-projection (via content prediction), thereby supporting a timely response initiation (see section above on how interlocutors anticipate turn-endings).

For example, Levinson and Torreira (2015) propose that prediction allows comprehenders to extract the speech act of the incoming utterance at the earliest possible time point, which will then allow some time for advance planning while still listening (early planning initiation), and thus timely turn-taking. Magyari and de

Ruiter (2008) propose that higher predictability of the sentence-final-words allows for successful turn-end projection, and thus timely turn-taking (timely response initiation). That is, when interlocutors listen, constraining information in the earlier parts of their interlocutors' sentences might help to anticipate the upcoming content and thereby even the turn-end (see also, Magyari et al., 2014, 2017). Prediction has also been proposed to assist in comprehending the new information part of the sentence, in a framework that attempts to bring together the notions of information structure, superficial (good-enough) language processing and prediction (Ferreira and Lowder, 2016). To this end, while listening to the phrase containing the given information of the sentence, a set of predictions is generated, which will assist in comprehending the phrase that specifies the new information in the sentence. The faster the incoming utterance is comprehended the faster one can start preparing a response, thereby possibly allowing for a fast take-over of the turn.

But prediction might not only affect turn-taking-timing by allowing for a timely planning initiation or a timely response initiation. If we assume that the higher probability a prediction is assigned, the easier the comprehension should be, then one can argue that listening to predictable content may be easier than listening to unpredictable content. Effects of predictability on processing speed have indeed been demonstrated; whereby predictable words were processed faster than less predictable ones (see for example Altmann & Kamide, 1999; Kliegl, Nuthmann, & Engbert, 2006; Traxler & Foss, 2000). More predictable words may thus have a processing advantage over less predictable ones. Content prediction can be based on semantic, morphosyntactic and prosodic information (for a short overview on prediction in language comprehension see Pickering & Garrod, 2013). Depending on when in time such prediction might arise, it could offer a way of making listening easier and thus make additional resources available for production planning. This in turn could help in overcoming part of the interference that arises because of the production-related processes and result not only in faster turn-taking –timing, but also in better listening and/or production.

That is, rather than focusing on how prediction might assist in extracting the speech act or anticipating the turn-end, I focus on how predictable speech content might improve listening quality, by making listening easier, and thereby decreasing the interference between listening and speech planning. This question is addressed in chapters 4 and 5.

Dependent Measures

Throughout the chapters, response latency (either verbal response or button press) and accuracy were used to evaluate several of the hypotheses. To address one of the major questions in this thesis, recognition memory performance was evaluated via a sensitivity measure of signal detection theory (the sensitivity signal detection index d', see Green & Swets, 1966; Macmillan & Creelman, 2005). This measure was used because it is a robust measure that depends on stimulus parameters and remains roughly invariant when response bias is manipulated. Rather than assessing online listening, recognition memory performance evaluates the outcome of consolidation processes during listening. For example, if having to engage in another task during listening, consolidation processes could be disrupted. Thus, the impact of allocating resources to that other task, on listening, could be quantified by evaluating recognition memory performance. Effects of dual tasking on consolidation or retrieval memory performance have been reported in a number of studies (e.g. Fernandes and Moscovitch, 2000). Moreover, memory performance has been used to quantify listening performance in studies on hearing impairment and acoustic challenge and have revealed that acoustically degraded words or syllables are more difficult to remember, despite being perceived correctly (e.g. Heinrich & Schneider, 2011; Heinrich, Schneider, & Craik, 2008; Surprenant, 1999).Thus, recognition memory performance can be a reliable offline measure of how attending to one task may impact on another task. Actual online listening was tested in chapter 5 of this thesis, using electroencephalography. Throughout the thesis chapters I will use the term "listening quality" to describe a wider spectrum of observations pertaining to both online listening quality (chapter 5) and to recognition memory performance (chapters 3 and 4).

In addition to that, more continuous measures of the participants' performance were collected for chapters 3, 4 and 5. In chapters 3 and 4 pupillometry was applied to monitor for cognitive effort as well as memory encoding and retrieval processes. Task-evoked pupillary responses (TEPR) are changes in pupil size that have been shown to reflect processing demands (e.g. Beatty & Lucero-Wagoner, 2000; Kahneman, 1973; Laeng, Sirois, & Gredeback, 2012) as well as memory encoding and retrieval processes (e.g. Goldinger and Papesh, 2012). In chapter 5 electroencephalography (EEG) and in particular the N400 component, allowed for online monitoring of semantic processing quality, while planning a response.

Thesis outline

Chapter 2 addressed the issue of planning onset timing and planning efficiency as possible factors contributing to the timing of turn-taking. To this end, participants of Experiment 1 heard the same constant sentence, to which they had to respond by describing one of the two displays (left or right) that appeared on screen. To manipulate speech planning onset timing one version of this auditory sentence (early timing condition) had the critical information (left or right) early in the sentence, while the other version (late timing) had it at the end of the sentence. To manipulate planning efficiency, and in particular conceptual processing, the pictures of the displays were presented either upright or upside-down. Experiment 2 had the exact same setup as Experiment 1 only now participants had to produce an initial coordinate noun phrase rather than a simple noun phrase (Experiment 1). It was hypothesized that participants would start planning their response faster, when given the opportunity to retrieve the answer earlier (early timing condition) as compared to later (late timing) during listening. Moreover, it was hypothesized that making speech production more effortful negatively affects response speed. In particular, when participants start planning late in time (late timing), a more demanding production planning (upside-down pictures) results in slower responding compared to a less demanding production planning (upright pictures).

In chapter 3 the focus shifted to the listening side. Given the possibility that interlocutors plan while still listening, it was examined what the consequences of concurrent planning were on listening. In an exposure phase participants either engaged in word planning (planning task) or did not do so (no-planning task) while listening to single words. A recognition memory test followed the exposure phase. It was hypothesized that participants would be better at recognizing words heard in the no-planning than in the planning task, demonstrating that the cognitive effort of concurrent speech planning had affected spoken word processing. Moreover, monitoring of pupil dilation in the exposure phase reflects the cognitive effort required for speech planning. Finally, pupil dilation during the no-planning task of the exposure phase provides information on the efficiency of memory encoding.

CHAPTER 1: INTRODUCTION

How listeners cope with the dual demands of planning single words while listening to whole sentences (rather than single words, see chapter 3) was the topic of chapter 4. In an exposure phase participants either engaged in word planning (planning task) or did not do so (no-planning task), while listening to sentences. In addition, predictability of the sentence-final-words was manipulated by embedding these either in a constraining or in a non-constraining sentence frame. Listening performance was assessed through a subsequent recognition memory test that evaluated how well participants remembered the sentence-final-words of the exposure phase. It was hypothesized that participants would be better at recognizing sentence-final-words heard in the no-planning than in the planning task, demonstrating that the cognitive effort of concurrent speech planning had affected sentence-final-word processing. Moreover, it was hypothesized that listening to predictable sentence-final-words would alleviate part of the observed interference by concurrent planning. As a result, predictable sentence-final-words of the planning task would be remembered more often than unpredictable ones. As in chapter 3, monitoring of pupil dilation in the exposure phase reflects the cognitive effort required for speech planning. A control experiment (Experiment 2) identical to Experiment 1, but involving no concurrent planning, was conducted to confirm that any findings of Experiment 1 (and of chapter 3) were not dependent on differences in effort of picture-recognition-processes between the planning and no-planning task, but rather indeed on differences in planning effort.

In chapter 5 the focus switched from monitoring listening performance offline, to also monitoring listening performance online with electroencephalography (EEG). In this context the N400 was of particular interest. Again participants either engaged in word planning (planning task) or did not do so (no-planning task) while listening to sentences that ended either in an expected or in a semantically anomalous sentence-final-word (exposure phase). A recognition-memory test evaluated how well participants remembered the sentence-final-words of the exposure phase. It was hypothesized that the N400 semantic anomaly-effect for items heard during speech planning would be smaller than for items heard during no speech planning. Recognition memory was again expected to be poorer for items heard during planning as compared to items heard during no planning. Finally, expected sentence-final-words of the planning task would be remembered more often than semantically anomalous ones. Chapter 6 summarizes and discusses the findings.

Chapter 2

Speech planning timing affects turn-taking timing

Gerakaki, S., Sjerps, M.J. Speech planning timing affects turn-taking timing

Abstract

Turn-taking timing is thought to be fast, resulting in smooth transitions of turns. Anticipation of a turn-end has been identified as a major factor contributing to fast turn-taking. In this chapter we argue that turn-taking timing depends not only on turn-end anticipation but also on speech planning onset timing and partly on speech planning efficiency. In Experiment 1 participants heard the same constant sentence, to which they had to respond by describing one of two displays (left or right) that appeared on screen. To manipulate speech planning onset timing two versions of this sentence were used: In the early timing version the critical information (here in italic font) was presented early in the sentence (En het plaatje *rechts/links*, kan je beschrijven wat daar gebeurt? / On the display on the *right/left*, can you describe what is happening (there)?), while in the late timing version it was presented at the end of the sentence (En kan je beschrijven wat er gebeurt op het plaatje *rechts/links*? / And can you describe what is happening on the display on the *right/left*?). To manipulate planning efficiency, and in particular conceptual processing, the pictures of the displays were presented either upright or upside-down. Experiment 2 had the exact same setup as Experiment 1 only now participants had to produce an initial coordinate noun phrase rather than a simple noun phrase (Experiment 1). Both experiments provide support for the idea that turn-taking timing is affected by the amount of speech planned ahead of the turn-end of the other interlocutor. In particular, one component of planning ahead in a conversational setting -namely planning onset

timing- stands out as a major factor contributing to the gap duration, while a second component - planning efficiency - provided less conclusive results on its contribution. The results demonstrate that when given the opportunity, interlocutors will opt for planning ahead while listening.

CHAPTER 2: SPEECH PLANNING TIMING AFFECTS TURN-TAKING TIMING

Introduction

Informal every-day observations as well as a number of conversation-corpora studies suggest that people are very good at the coordination of turns (e.g. Sacks et al., 1974). For example, Stivers and colleagues (2009) observed that the overall cross-linguistic median for turn-taking timing in their study was a gap of 100 ms (see also Beattie & Barnard, 1979; Bosch, Oostdijk, & Boves, 2005). This line of research resulted in the idea that turn-taking timing is smooth (e.g. Garrod & Pickering, 2004; Pickering & Garrod, 2004; Scott et al., 2009; Stivers et al., 2009). Smooth turn-taking timing has often been described as the result of effectively anticipating or projecting a turn's end (Magyari & de Ruiter, 2008; Sacks et al., 1974; de Ruiter et al., 2006). Optimal performance in this anticipation process can lead to minimal gaps between the turns of the two interlocutors.

The current chapter proposes that inter-turn-intervals don't only depend on anticipatory mechanisms but also on the processing demands resulting from the need to listen and prepare a response. In particular we aimed at 1) investigating whether interlocutors will start planning their response to a question faster, when being able to retrieve the answer earlier as compared to later during listening; 2) demonstrating that raising the cognitive demand on the production side by manipulating planning efficiency might slow down turn-taking timing.

On turn-end projection and cues signaling turn-endings.

On describing the features of turn-constructional units Levelt (1993) draws on the theory by Sacks and colleagues (1974) and posits that a turn consists of one or more units which are projectable in the sense that each projectable unit can be roughly predicted from its type. As a result an interlocutor can know where to take over (Levelt, 1993). For example, the sentential unit *Could you tell me the way to...* could be completed by projecting into a noun phrase (*Stephen's Church*), Other possible projectable units of a given sentential unit would be clausal, phrasal and lexical.

The *projection theory* put forward by Sacks and colleagues (1974) involved anticipation or projection of turn end via structural and contextual information. This anticipation or projection mechanism allowed the next speaker to start talking at the projected turn ending. Around the same time another group of theories was formulated postulating that the next speaker will only start talking prompted by perceiving a signal that the turn of the current speaker is over, or will soon be over. This group of theories was labeled the *reaction or signal theory* (see for example Duncan, 1972). These two research lines have given an insight into the different cues that play a role in turn-end projection or signaling. Rather than discussing further the possible merits of each view, we will briefly review the cues that have been implicated in turn-taking coordination.

The existing literature on turn-taking timing has unveiled a large number of cues that are useful to the listener in predicting when a turn is going to end (see also Levinson & Torreira, 2015). Cues can be derived from face-to-face communication, like information from eye-gaze (e.g. Duncan & Fiske, 1977; Kendon, 1967). But cues can also come from the speech signal. Syntax was identified as the main source for projecting turn-ending points by Sacks and colleagues (1974). Moreover research has pointed to pragmatic information (Ford & Thompson, 1996) as well as prosody (Beattie, Cutler, & Pearson, 1982; Couper-Kuhlen & Selting, 1996; Cutler & Pearson, 1986; Local, Kelly, & Wells, 1986). For example, a high pitch peak can signal that the turn will end at the next syntactic completion point (Schegloff, 1996). Such a signal would allow for ample time to anticipate a turn’s ending and maybe even start planning the reply. In an analysis of conversational corpora Ford and Thompson

(1996) identified so called complex transition relevant places (TRP). These were points in the conversation, in which all three types of cues (syntactic, intonational, and pragmatic) coincided. The authors found that nearly three quarters of all speaker changes occurred at a complex TRP. Such a finding might point to the fact that the cues used have an additive effect. The more cues point towards a turn-ending, the more salient this becomes.

In a somewhat more experimental approach Schaffer (1983) used passages from conversations and applied low-pass filtering to create stimuli with unintelligible speech and intact intonation. She concluded that lexicosyntactic information, compared to intonation, offers a more reliable source to anticipate a turn's ending. In an attempt to further clarify the relative contributions of lexical and syntactic information and intonation to turn-end projection de Ruiter and colleagues (2006) had participants listen to turns from natural conversations and press a button when they thought the turn they were listening to ended. In the stimuli they manipulated the presence of lexical and syntactic information as opposed to intonational contour. They found that lexical and syntactic information was a reliable cue in predicting turn endings, while the intonational contour seemed not to offer adequate information (but see Bögels & Torreira, 2015). Magyari and de Ruiter (2008) further linked this anticipatory advantage of lexico-syntactic information to the higher predictability of the last words in the utterance. In their study they had participants listen to the same turns from natural conversation as in de Ruiter and colleagues (2006). Subjects made better predictions about the last words of those turns that had more accurate responses in the earlier button press experiment. To conclude, research on turn-end projection has offered reliable evidence that turn-end anticipation is one major factor affecting turn-taking timing.

Planning an utterance. Some factors affecting the scope of incremental planning.

During conversation an interlocutor not only has to pick up cues that will signal the end of a turn. She also has to comprehend the input and at some point in time start planning her reply. In order to get closer to the core of conversational interaction, it would be useful to include the listening-production component in the equation describing the timing of turn-taking. The essence of regulating turn-taking is exactly this interplay between listening and production.

Namely, to produce an utterance one needs to start planning and to start planning one needs to decide what message one wishes to convey. To decide what message one wishes to convey, one usually has to first comprehend the incoming utterance. For example, Pope and colleagues (1971) and Siegman and Pope (1972) have offered one of the few indications that cognitive load resulting from the interviewer's ambiguous questions, does affect pause and switch duration of the interviewed, making them longer. This somewhat naturalistic observation presumably signifies that the person being interviewed needed more time to conceptually prepare and thus also more time to plan her reply.

In this section we focus on the role of production planning and argue that turn-taking timing depends also on characteristics of the to be prepared speech signal. That is, at some point in time the interlocutor has to convert the message she wants to convey into speech signal. This process requires going through a number of processing levels (see for example Bock, & Levelt, 1994) including computing the conceptual, grammatical, phonological and phonetic representations prior to initiating articulation. Thus, the speed with which these computations can take place is also of essence for the resulting response timing.

In addition, it is widely accepted in the literature that speech is planned incrementally. Incrementality is the idea that people plan upcoming parts of their utterance while articulating earlier parts of it. Even though incrementality is a generally accepted idea in the field, there is no consensus on the size of the incrementally planned unit. The size of the incrementally planned unit is often called the scope of incremental planning and has been shown to be sensitive to manipulations of time pressure (Ferreira and Swets, 2002), cognitive load (Wagner, Jescheniak, & Schriefers, 2010) and working memory load (Slevc, 2011). Thus, in addressing the question of how speech planning might affect turn-taking timing in conversation, one has to take into account that the scope of incremental planning is rather flexible (Ferreira & Swets, 2002; Slevc, 2011; Wagner et al., 2010) and that there is still no consensus on the size and content of the planning units at different levels of processing (Wheeldon, 2012). Despite that, research looking into the scope of incremental planning offers some indications that the complexity of the to be produced utterance and the time available to prepare speech could affect response timing.

Evidence that the complexity of the utterance might affect the timing of speech production comes from research that used the so-called online picture description task to investigate the planning scope in sentence production (Martin, Crowther, Knight, Tamborello, & Yang, 2010; Smith & Wheeldon, 1999). Participants are asked to describe the movement of multiple objects on screen. The critical manipulation concerns the syntactic structure of the planned and produced sentences. For example, the movement of the objects might trigger starting with a simple or a complex noun phrase. Such a setup tapped into grammatical encoding processes, namely the need to select the appropriate lexical items and to generate a syntactic order that will allow positioning these lexical items in a linear order (Wheeldon, 2012). Smith and Wheeldon (1999) found that planning initial complex noun phrases results in significantly longer production onset latencies than planning initial simple noun phrases. They interpreted this finding as evidence for complete lemma access for the first phrase of the utterance, and only higher level conceptual processing of the remainder of the utterance. That is, the scope of incremental planning was always the first noun phrase. But an initial complex noun phrase resulted in longer response latencies compared to an initial simple noun phrase.

Evidence that the time available to prepare speech might affect the timing of speech production comes from research investigating the effect of prosodic structure on production latencies in prepared and on-line speech (Wheeldon and Lahiri, 1997). Prepared speech resembles more the conversational situation in which an interlocutor has already prepared a response but has to buffer it until the other interlocutor's turn has finished, while on-line speech is more similar to a situation in which speech is uttered immediately the moment parts of it are ready for production. Wheeldon and Lahiri (1997) concluded that response time in prepared speech production depends on the number of phonological words a sentence contains, whereas response time in on-line speech production depends on the complexity of the to be produced phonological word. That is, whereas complexity of the phonological word plays a role in on-line speech, it does not seem to contribute when uttering prepared speech. Thus, not only differing processing demands due to the need to produce utterances of different complexity (e.g. initial complex as opposed to simple noun phrases) can trigger different response timing, but also on-line and prepared speech seem to invoke different processing strategies that can result in differences in response timing. Importantly, producing a rather simple utterance might not be affected by when in time planning could initiate, while producing a more complex utterance might benefit from an early planning onset.

Beyond turn-end projection. The current study.

To summarize, we propose that while the reviewed anticipatory tools are an important factor contributing to turn-taking timing, much of the variation in turn-taking timing might also be captured by the listening-production interplay demands. Having an estimate of a turn's end is definitely of advantage, but it is not the only prerequisite in initiating a timely response. An interlocutor might have a good estimate of a turn's end, but still have no reply planned. Thus, when the time comes for her to speak, there is still nothing to produce. Whether there is something ready to produce in time depends on 1) how complex the to be comprehended input stream was (we call this listening efficiency), 2) how complex the to be produced output stream was (we call this efficiency of planning) but also 3) on the time available to plan (we call this planning onset timing). In this chapter we focus only on planning onset timing and efficiency of planning and not on listening efficiency.

The first aim of the current chapter was to investigate whether -when given the opportunity- interlocutors will start planning their replies before the offset of their interlocutors turn, thereby resulting in faster response timing, as opposed to only when that turn is finished. To answer this question we manipulated the actual time-point during which speech planning could initiate. The second aim of the current chapter was to investigate in how far varying the cognitive effort involved in producing speech (planning efficiency) would affect the size of the inter-turn-interval, with the need to produce a complex output (low planning efficiency) resulting in a larger inter-turn-interval compared to the need to produce a simpler output (high planning efficiency). To answer this question we manipulated the complexity of the to be produced utterance by presenting the to be named pictures either upright or upside-down, thereby affecting conceptual processing.

To conclude, the proportion of speech planned while listening to the other interlocutor might differ depending on a) when this planning started and b) how efficient the planning was. To control the effect of anticipation and allow for effects of planning onset timing and planning efficiency to be observed, we used a constant predictable structure for the input sentences. This decision admittedly maximized anticipatory performance and made listening very easy, but by keeping these two factors constant we could more clearly assess what the contribution of planning onset timing and planning efficiency to turn-taking timing is (unconstrained from any possible interactions with listening complexity). Moreover, the need to have some control over the size of the planning unit led to the decision to use a variant of the paradigm that was used by Smith and Wheeldon (1999) and Martin and colleagues (2010) to investigate the planning scope in sentence production.

Experiment 1

Method

Overview. In the first experiment we investigated whether planning ahead of an utterance is a parameter that can affect turn-taking timing in a conversational setting. In particular, we were interested to see how the time point of planning onset, and consequently the amount of information planned ahead, might affect the speed with which an individual is able to take over a turn. If one has sufficient time to plan ahead an utterance, then she will have an almost ready utterance to produce when the time comes to take over a turn. If though it is not possible to plan ahead an utterance, then the whole planning phase, or at least some elements of it, will have to go into the turn-transition gap.

In addition, we manipulated planning efficiency by varying the difficulty of conceptual processing. To this end, we included a condition in which the target pictures were presented upside-down. If planning ahead is slowed down by difficulty in conceptual processing, and this slow-down also results in larger inter-turn intervals, then indeed planning efficiency affects turn-taking timing.

Importantly, we set up an experimental environment that involves both listening and speaking for the participant, while offering a controlled setting. Even though this setting is far from natural conversation, it still includes two of the most basic parameters in conversation; namely, listening and production combined in one study.

Participants. 20 people from the Max Planck Institute for Psycholinguistics subject pool participated in the study. All were native speakers of Dutch and had normal or corrected-to-normal vision. None of them reported a speech or hearing problem, and none had been diagnosed with dyslexia.

Design and Materials. One hundred and eighty (180) pictures were used to create thirty displays. The pictures were selected out of four databases (Bates et al., 2003; Severens et al., 2005; Snodgrass & Vanderwart, 1980; MPI picture database; see Appendix A for picture-names). The majority of pictures (134) were selected from Severen's database, with the other databases only serving as an additional source when necessary. A total of eleven additional new pictures were created for this experiment. Each display consisted of six phonologically and semantically unrelated pictures. A solid black line separated the screen into a left (L) and a right (R) part. Three out of the six pictures of a given display were positioned on the left part of the screen and the remaining three pictures were placed on the right part of the screen. The pictures on each side (L, R) were grouped in an A+BC manner, meaning that the first picture stood alone on the left side of the arrow, while the remaining two items were grouped together right of the arrow. This setup was used to make participants produce initial simple Noun Phrases (NP) for all responses. *Figure 1* depicts an example display from the practice material.

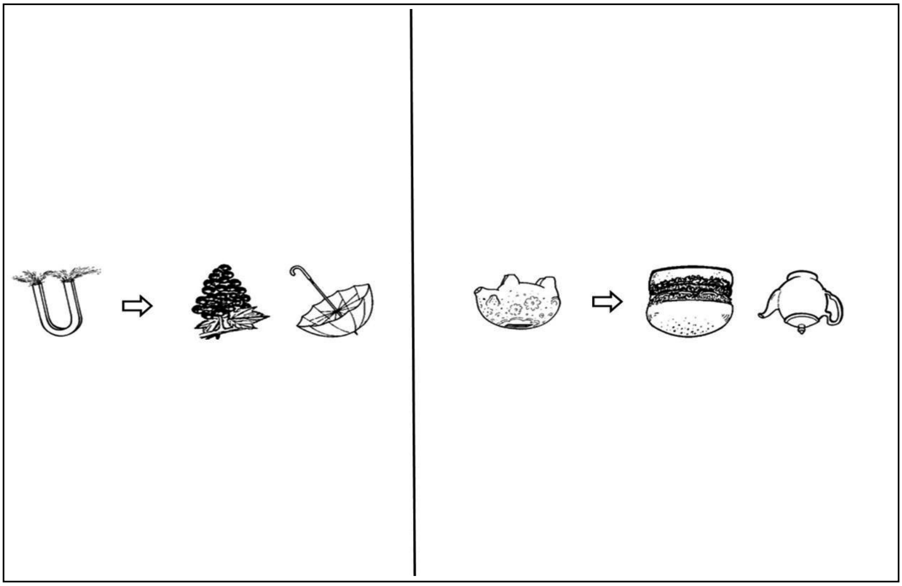

Figure 1. Example of one upside-down display used in the practice session of Experiment 1. The vertical line splits the screen in two parts. A left and a right part.

Participants saw each display in the following four conditions. Early timing with upright orientation of the pictures (EU), early timing with upside-down orientation of the pictures (EUD), late timing with upright orientation of the pictures (LU) and late timing with upside-down orientation of the pictures (LUD). All six pictures of a given display were presented upright when in the U conditions and upside-down when in the UD conditions.

Two auditory stimuli were combined with each of these thirty displays to create the four conditions of the experiment (early upright, early upside-down, late upright, late upside-down), resulting in 120 trials per participant. The purpose of these auditory stimuli was to inform the participant which display she should describe in this given trial (“left” or “right”). Thus each auditory stimulus had a "left" and a "right" version. This was necessary to create uncertainty about which part of the screen one would have to describe next, thus only allowing speech planning to initiate once the participant knows which side of the screen she should describe (left or right; see *Table 1* for an overview of the conditions).

All auditory stimuli were processed in such a way as to have equal length (3217 ms). In the early left/right condition the duration from the offset of “left/right” to the offset of the auditory stimulus as a whole, was approximately 1982 ms. Thus, if participants wanted to start planning ahead their response (while still listening), then they had a maximum of 1982 ms to do so. In the late left/right condition the time available to plan ahead was zero, given that “left/right” was the final word of the auditory stimulus. At best, participants could start differentiating between hearing “left” or “right at the onset of the first consonant of these words, resulting in some 560 ms available time until auditory stimulus offset. Note that by analogy, the time available for planning ahead in the early condition would now be 2487 ms, if rather than setting the offset of “left” / ”right” as the initial timepoint during which planning ahead could start, the onset of the first consonant of “left” / ”right” was set as the initial timepoint during which planning ahead could start, instead.

The one hundred and twenty (120) trials were split into four blocks of thirty trials each (one display equals one trial). In each block all thirty displays were presented once. Overall, each of the thirty displays was presented once in each condition for each participant, resulting in four repetitions of the exact same display within participant (one presentation per block). To avoid any condition-order effects, the order of appearance of each condition was counterbalanced across participants for each display and for blocks, and also within participants across displays.

Table 1. The conditions of Experiment 1. The critical information of the auditory stimuli could specify to describe either the left or the right part of the display. English translation: E: "On the picture on the right/left, can you describe what happens there?" L: "And can you describe what happens on the picture on the right/left?".

Condition	**Timing of critical information**	**Orientation**	**Auditory stimulus**
EU	Early	upright	En het plaatje **rechts/links**, kan je beschrijven wat daar gebeurt?
EUD		upside-down	En het plaatje **rechts/links**, kan je beschrijven wat daar gebeurt?
LU	Late	upright	En kan je beschrijven wat er gebeurt op het plaatje **rechts/links**?
LUD		upside-down	En kan je beschrijven wat er gebeurt op het plaatje **rechts/links**?

The experiment was a 2X2 (early/late X upright/upside-down) within subjects design. Dependent measures consisted of timing (gap duration) and accuracy of the response. The early upright (EU) and early upside-down (EUD) conditions revealed the critical information (left or right) early on in the auditory stimulus, while the late upright (LU) and late upside-down (LUD) conditions only revealed the critical information close to the offset of the auditory stimulus.

To avoid that participants through display repetition learn to anticipate which side of the screen they would be asked to describe for a given display, out of the four namings of a display by a participant, two were about the left part of the screen and two about the right part. This left-right naming was also carefully counterbalanced across conditions and displays for all participants, so that all participants had equal data points per condition and side (R or L). Extra care was taken so that the pictures of the left side of a given display approximately matched the characteristics of the pictures on the right side of that same display. Four experimental lists were used in this experiment.

Procedure. The experiment took place in a sound treated booth at the Max Planck Institute for Psycholinguistics. Participants received monetary compensation for their participation. The experiment lasted approximately one hour. The participant was seated in front of a computer screen wearing headphones and was asked to listen to what is being asked of him and reply to that as quickly as possible. The task was an adaptation of the online picture description task by Smith & Wheeldon (1999).

Presentation software was used for controlling stimuli presentation (Neurobehavioral Systems, Albany, CA). A trial started with a blank screen (800 ms), followed by a display like in *Figure 1*, while at the same time the auditory stimulus was heard. The display stayed on screen for a maximum of 8000 ms, after which the next trial automatically started, or until the participant was almost done uttering her response, and thus the experimenter pressed the enter key to proceed to the next trial. This was meant to allow the experimenter to adapt the pace of the experiment to each individual participant, so that sufficient pressure was exerted to provide fast responses.

For example, the display in *Figure 1* should elicit the response (target on right side) *“Het varken gaat naar de hamburger en de theepot”* (English translation: “The pig goes to the hamburger and the teapot”). Participants were instructed to “translate” the arrow into “gaat” (English translation: "goes"). A microphone recorded response content (maximum duration of the recording was 4000 ms), so that RT s could later be computed through annotation in Praat (Boersma, and Weenink, 2012, version 5.1).

After participating in the main experimental task, some additional measures were collected. To evaluate any relation between turn-taking timing performance and verbal ability, a questionnaire on speaking experience was also administered to the participants, but this is not further evaluated in this chapter.

Hypotheses. If turn-taking timing is affected by planning onset timing, then having the opportunity to start planning early will lead to a smaller inter-turn interval than when only having the opportunity to start planning late. If the effect of planning on turn-taking timing is not only dependent on the time-point at which one could start planning her response, but also on planning efficiency, then participants should be faster to take over a turn when conceptual processing is easy (U conditions), than when conceptual processing is hard (I conditions). This orientation effect (U vs I) should interact with timing of planning onset (E vs L), with the orientation effect being larger for late than for early timing. In particular, the slow-down by upside-down picture presentation as compared to upright presentation should be biggest in the late condition, when people don't have time to plan while listening.

Results

Accuracy. All recordings were annotated for response time and errors using the Praat software. Errors (10.65%) were categorized into the following categories. *Naming errors* consisted of cases in which participants used another word to refer to one of the target pictures, mispronounced one of the words or used the wrong article for the given object. *Timeout errors* refer to cases in which the speaker did not manage to complete her description in time.

The accuracy data were analyzed using generalized linear mixed effects regression models in R (version 2.14.2; The R foundation for statistical computing; lme4 package; Bates, Mächler, Bolker, & Walker, 2015). The models were fitted with binomial distributions (Jaeger, 2008). All factors justified by the design of the study were included in the initial model as fixed structure, while the random structure of the models was determined following the suggestions of Barr, Levy, Scheepers, and Tily (2013). Using a backward elimination procedure any interaction that did not improve model fit was excluded from the model. But all predictors justified by the design were kept in the models. In addition, a null model that included only the random structure of the models was established. Model comparison was made using log-likelihood ratio tests. Statistics of the optimal models are reported.

Error rates were evaluated with a model including the factors timing (early and late; fitted as -1 and 1, respectively), orientation (upright and upside-down, fitted as 1 and -1 respectively) and experimental list (centered), as well as the interaction of timing with orientation as fixed effects, and intercepts by participant and display, with slopes for timing and orientation by participants as the random effects[1]. None of the factors contributed significantly to predicting error rates (intercept = -2.3729; timing by orientation: b = -.0574, p =.407; timing: b =.088, p =.260; orientation: b = -.1017, p =.215; list: b = -.081, p =.566).

Response Time. Incorrect responses (10.65%) as well as data points with a value further than two standard deviations away from the participant mean (4.71%) were removed from the data. Response Times (RTs) were measured from the offset of the auditory stimulus. *Table 2* summarizes the Response Time (RT) data for the four conditions. The same model as for accuracy was run for response time. Only the factor timing (intercept = 435.54, b = 68.57, p <.001) contributed significantly to predicting response time (orientation: b = -4.37, p =.298; list: b = 7.15, p =.745).

Table 2. Mean RTs (in ms) in the various experimental conditions of Experiment 1. RTs were measured from the offset of the auditory stimulus.

Timing	Upright Orientation	Upside-down Orientation
early	363	370
late	500	507

Discussion

The findings of Experiment 1 confirmed our hypothesis that turn-taking timing can be affected by planning processes. When planning could initiate early, participants made use of the time they had to plan. This resulted in significantly smaller gaps in the early than in the late condition. This indicates that in the early timing condition participants had possibly initiated some aspects of the planning process before the offset of the turn. Given that in the early timing condition participants had about 1982 ms time from the offset of the words “left” or “right” until the offset of the sentence as a whole (see methods section) there was ample time to prepare. That time was though still not too long, since there were only very few cases of overlap between listening and speaking. However we did not observe any effect of our manipulation of planning efficiency (upright vs. upside-down picture presentation) on turn-taking timing. Perhaps preparing to name a simple noun phrase, followed by a complex noun phrase is not demanding enough, even when the pictures are presented upside-down, because participants could prepare the name of the first picture fast and then opt to incrementally prepare the rest of the utterance (complex noun phrase) while already producing the first part of the utterance (simple noun phrase).

Experiment 2

Method

Overview. In Experiment 1 we provided evidence that participants start planning their utterance before the offset of their interlocutor's turn, when given the opportunity. At the same time we did not find any evidence that planning efficiency, as measured via conceptual processing speed of upside-down pictures, did affect turn-taking timing. Possibly the fact that the to be produced utterances started with a simple noun phrase (NP), allowed for high planning efficiency in both upright and upside-down displays, because one could start planning incrementally the first simple noun phrase, while also preparing the second complex noun phrase. Therefore in Experiment 2 we chose to invert the order of the noun phrases. Now the initial noun phrase was a complex noun phrase, while the second one was a simple noun phrase. According to Smith and Wheeldon (1999), complex initial NPs lead to slower response times than simple initial NPs. If this is true, then we might be able to manipulate planning efficiency more successfully, thereby providing some evidence on the role of planning efficiency in turn-taking timing.

Participants. 22 people from the Max Planck Institute for Psycholinguistics-subject pool participated in the study. All were native speakers of Dutch and had normal or corrected-to-normal vision. None of them reported a speech or hearing problem, and none had been diagnosed with dyslexia.

Materials. The exact same displays and conditions as in Experiment 1 were used. Only now the to be described pictures started with a complex NP (A & B go to C) rather than with a simple NP (A goes to B & C). To this end, the pictures on each side (L, R) were grouped in an AB+C manner, meaning that two of the three items were grouped together left of the arrow, while the third picture stood alone on the right side of the arrow. This setup was used to make participants produce initial coordinate Noun Phrases (NP) for all responses. Figure 2 depicts an example display from the practice material.

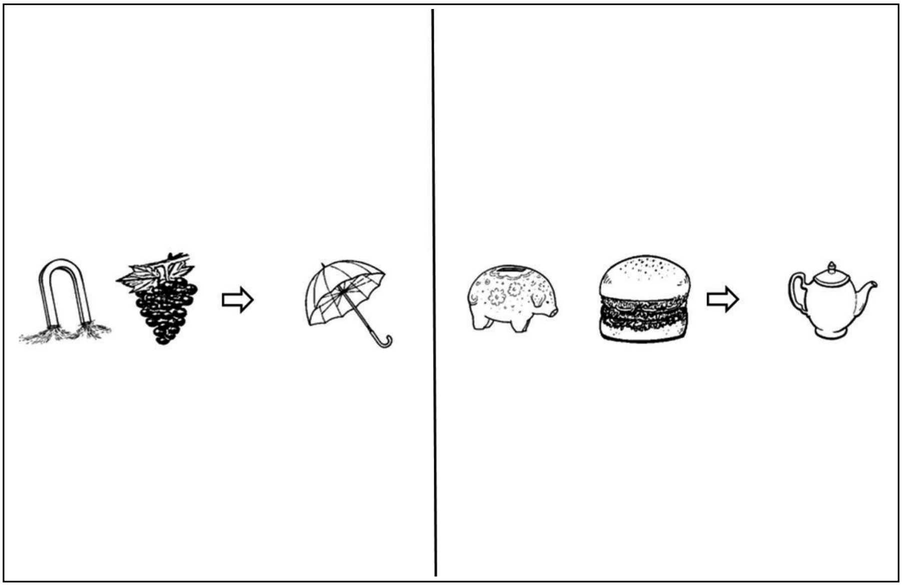

Figure 2. Example of one upright display used in the practice session of Experiment 2. The vertical line splits the screen in two parts. A left and a right part.

Procedure. The procedure was identical to Experiment 1.

Hypotheses. In Experiment 2 we expect to find the same pattern of results as in Experiment 1 in terms of onset time of planning; only now we expect the orientation effect as well as its interaction with planning onset time (see hypotheses of Experiment 1 for more details) to stand out more clearly, since now the initial noun phrase was complex rather than simple.

Results

Accuracy. All recordings were annotated for response time and errors using the Praat software. Errors (10.64%) were again categorized into *Naming errors* and *Timeout errors*. The accuracy data were analyzed using generalized linear mixed effects models. The same model as in Experiment 1 was tested: Timing (early and late; fitted as -1 and 1, respectively), orientation (upright and upside-down, fitted as 1 and -1 respectively) and experimental list (centered), as well as the interaction of timing with orientation were the fixed effects, and intercepts by participant and display, with slopes for timing and orientation by participants as the random effects. None of the factors contributed significantly to predicting error rates (intercept = -2.3714; timing by orientation: b = -.0217, p =.747; timing: b =.0001, p =.999; orientation: b = -.1408, p =.089; list: b =.0282, p =.833).

Response Time. Incorrect responses (10.64%) as well as data points with a value further than two standard deviations away from the participant mean (4.48%) were removed from the data. Response Times (RTs) were measured from the offset of the auditory stimulus. *Table 3* summarizes the Response Time (RT) data for the four conditions. The same model as for accuracy was run for response time. Timing (intercept = 429.80, b = 82.32, p <.001) as well as the interaction of timing with orientation (b = -8.13, p =.043) contributed significantly to predicting response time, while orientation (b = -5.871, p =.166) and list (list: b = -1.174, p =.968) did not.

Separate analyses were carried out for the late and early planning conditions to evaluate the interaction. Orientation (upright and upside-down, fitted as 1 and -1 respectively) and list (centered) were included as fixed factors, with intercepts by participant and display, and slopes for orientation by participants as the random effects. The analysis revealed that participants were significantly slower to respond when the pictures were upside-down compared to in upright position, but only in the late timing condition (intercept = 513.36, b = -15.032, p =.041) and not in the early (intercept = 346.88, b = 2.844, p =.6291).

Table 3. Mean RTs (in ms) in the experimental conditions of Experiment 2. RTs were measured from the offset of the auditory stimulus.

Timing	Upright Orientation	Upside-down Orientation
early	353	343
late	491	518

Discussion

Experiment 2, like Experiment 1, confirmed our hypothesis that turn-taking timing can be affected by speech planning onset timing. But in addition to that we observed that turn-taking timing can also be affected by planning efficiency: when planning could initiate early, participants made use of the time by preparing the picture names in time and avoiding any effect of orientation (high planning efficiency). This was not the case when planning could only initiate late. There participants could not prepare the names of the upside-down pictures as fast as they did for the upright pictures. As a result, they were slower to initiate a response when having to describe upside-down pictures (low planning efficiency). Thus planning efficiency, even in the quite basic form as tested in the current experiment, affected turn-taking timing.

General Discussion

In two experiments we investigated the relationship between planning onset timing and planning efficiency, and turn-taking timing. Response speed data from both experiments support the hypothesis that planning onset timing is an important factor contributing to turn-taking timing. Having the opportunity to start planning a reply early (i.e., while listening to the other interlocutor) results in a shorter turn transition. In other words, preparing speech while listening, may play a role in achieving smooth conversation. In recent years, further research has provided additional support that interlocutors start planning their turn at least shortly before the turn-end of their interlocutor (Bögels et al., 2018, 2015; Boiteau et al., 2014; Sjerps & Meyer, 2015).

Evidence for the role of planning efficiency on turn-taking timing was less strong. In Experiment 1, where participants had to prepare an utterance starting with an initial simple noun phrase followed by a complex noun phrase, no effect of planning efficiency (orientation) on turn-taking timing was observed. However, when the initial noun phrase was complex (Experiment 2), planning efficiency did interact with planning onset timing. As a result, when planning could only initiate late in time, participants were affected by the orientation of the pictures: they needed more time to initiate a response when the pictures were upside-down as compared to upright. This delay was not observed when planning could initiate early in time, suggesting that participants used part of their interlocutor's response window to prepare their utterance, and compensate for any increase in planning difficulty.

The fact that planning efficiency interacted with planning onset timing when having to prepare an initially complex noun phrase (Experiment 2) but not when having to prepare an initially simple one (Experiment 1), may be linked to the discussion on planning scope (Ferreira & Swets, 2002; Slevc, 2011; Wagner et al., 2010). In Experiment 1, participants could in principle have followed the pattern observed in Smith and Wheeldon (1999) and have thus opted for preparing only the first part of the utterance (simple noun phrase) before initiating speech, and subsequently planning the rest of the utterance. This way the conceptual processing even of upside-down pictures could finish in time. In Experiment 2 again participants likely opted to prepare the first part of the utterance, but now the initial noun phrase was complex and when little time was available to prepare (late condition) planning efficiency (orientation) affected response time.

In the current experiments participants were under explicit time pressure, as displays disappeared from the screen as soon as the participant was about to utter the final word of her description (the disappearance of the display was controlled by the experimenter). This design may have induced participants to respond relatively fast thereby making the planning scope as minimal as possible. It was only when planning could initiate late, and the to be produced utterance was an initial complex noun phrase, that adhering to this principle became very difficult.

The early late manipulation introduced in the current experiments resembles in a way the prepared and on-line speech production manipulation of Wheeldon and Lahiri (1997). Similarly to their finding that complexity of the phonological word plays a role in on-line speech, while it does not seem to contribute when uttering prepared speech, we observed that planning efficiency (orientation) only played a role when participants did not have enough time to prepare (late condition of Experiment 2; in a way on-line speech) and not when they had sufficient time to prepare (early timing condition; in a way prepared speech). A study investigating the neural correlates of word production in immediate and delayed naming, concluded that delayed naming only captures encoding processes up to the beginning of phonological encoding, whereas immediate production captures the whole spectrum of processes involved in speech production (Laganaro & Perret, 2011). This might explain part of the added difficulty in preparing a response for immediate production.

Future research, implementing more demanding manipulations of planning efficiency, like introducing uncertainty on which syntactic structure to prepare (preferably in a within participants design), or possibly even rendering conceptual or lexical processing more difficult, might better illustrate its role in turn-taking timing. Moreover, rather than using different input sentences thereby making the input important to listen to, in the current experiments we kept the sentence constant. Future work should investigate how planning efficiency and thus turn-taking timing is affected by the need to actively listen to an unpredictable input that in addition needs to be monitored for comprehension and or correctness. Such a setting in which the need to allocate processing resources becomes evident, might allow for more insight on how comprehension and production interact and how this impacts on planning efficiency and turn-taking timing.

Conclusion

Anticipation of a turn's end has been proposed as the central factor governing the timing of turn-taking. Here we provide evidence for a richer interplay of factors. Following the results of our study we argue that turn-taking timing is not only the product of anticipation but also of planning onset timing and to some extent of planning efficiency.

Footnotes

[1] The random structure is described once for each dependent measure. Only if convergence problems forced us to alter it in any way it is mentioned again. When not explicitly stated, the random structure described at the beginning of each section applies to all the models of this dependent measure.

Chapter 3

Concurrent speech planning affects memory for heard words

Gerakaki, S., Sjerps, M.J. Concurrent speech planning affects memory for heard words.

Abstract

In conversation interlocutors can swiftly switch from listener to speaker role, suggesting that at times speech planning and listening overlap. In chapter 2 participants started planning their response as early as possible, even in overlap with listening (early-planning condition). However both speech production and perception may require a considerable amount of attention, which could cause interference between the two processes. In the current chapter we investigated how well listeners processed words they heard while planning other words. In an exposure phase participants either engaged in word planning (plan task) or did not do so (no-plan task) while listening to spoken words. Measurements of pupil dilation confirmed that cognitive effort was required for speech planning. In a subsequent recognition memory test participants were better at recognizing words heard in the no-plan than in the plan task, demonstrating that the cognitive effort of concurrent speech planning had affected spoken word processing. This effect was independent of incidental (Experiment 1, surprise memory test) or intentional (Experiment 2, anticipated memory test) encoding mode. Pupil dilation further linked the effect to memory encoding and retrieval processes. The need to share resources between speech planning and perception may constrain everyday turn-taking behavior.

Introduction

The continuous switching of speaker roles between interlocutors in conversation (also referred to as turn-taking) is said to be remarkably fast and smooth (e.g., de Ruiter et al., 2006; Gambi & Pickering, 2011; Pickering & Garrod, 2013; Scott et al., 2009; Wilson & Wilson, 2005). In Dutch, for example, Stivers and colleagues (2009) have estimated that the average gap between turns is 109 ms, and Ten Bosch, Oostdijk, and Boves (2005) reported a median gap of 330 ms. Similar inter-turn intervals have been reported for English (Beattie & Barnard, 1979: a median value of 333 ms). This is notable because even naming a picture with a simple noun has been shown to last at least 600 ms (Indefrey & Levelt, 2004; Indefrey, 2011), and preparing the first phrase of a sentence can easily take up more than a second (Konopka, 2011).

Taken together, then, the available estimates of turn-taking timing and those of the time needed for single word and sentence production suggest that in dialogue speakers often begin to plan their utterance while listening to the interlocutor. Indeed, recent research has shown that, at the very least, interlocutors start planning their turn shortly before the turn-end of their interlocutor (Bögels, Magyari, & Levinson, 2015; Boiteau, Malone, Peters, & Almor, 2014; Sjerps & Meyer, 2015). The findings in chapter 2 also indicated that, under some circumstances, participants started to plan their responses to the questions as early as possible. In situations of parallel listening and planning, cognitive resources may thus have to be shared between the two processes, which may have negative effects on the quality of speech processing. Alternatively, however, given the extensive practice people have in concurrent listening and speech planning, only minimal cost might arise. To understand the pressures that govern turn-taking behavior it is thus critical to establish to what extent dual-task interference between listening and planning may play a role. As a first step in addressing this question we investigated the interplay between perception and production in a controlled task environment, examining how well people encode single words that they hear while they are planning speech.

Cognitive interference in listening and speaking.

Cognitive interference between speech planning and speech perception may come about as a result of two types of processing overlap. First, perception and production may involve language-specific processes with overlapping or competing representations. To exemplify, it has been shown that the presentation of written or spoken words can affect single word production due to semantic or phonological relatedness (e.g. Damian & Martin, 1999; Schriefers, Meyer, & Levelt, 1990). While such influences are bound to play an important role in turn taking situations, and may even contribute to alignment (Garrod & Pickering, 2004), the current research focused on a second type of interference; their shared reliance on domain-general central processing resources.

Evidence demonstrating that both speaking and listening require central processing resources (and may, hence, interfere with each other) comes from studies that used dual-task designs in which one task was linguistic (speech perception or speech production) and the other non-linguistic. For instance, Kubose and colleagues (2006) used a driving simulator to test how driving performance was affected by producing or comprehending speech. They found that both production and comprehension had a strong impact on driving performance (see also Strayer & Johnston, 2001). The reverse is true as well. Driving reduces language production and comprehension abilities as evidenced by poorer story retelling performance and a negative impact on story comprehension and long-term memory encoding (Becic et al., 2010).

In addition to those sentence-level demonstrations of central interference, single word perception (e.g. Cleland, Tamminen, Quinlan, & Gaskell, 2012), individual phonemes (e.g.

Gordon, Eberhardt, & Rueckl, 1993; Mattys & Wiget, 2011) and sentences (Bosker et al., 2017) have been shown to be affected by concurrent tasks (see also Mattys and colleagues, 2009; on how lower level processes of speech perception are affected by domain-general resource depletion).

Single word production planning has also been shown to absorb general cognitive resources. For example, Ferreira and Pashler (2002) demonstrated in a dual-task study that response speed in picture naming and in a concurrent secondary nonlinguistic task (tone-discrimination) were modulated by the ease of lemma and word form selection. Cook and Meyer (2008) reported evidence of central capacity demands even for phonological processing (see Roelofs & Piai, 2011, for a review on attention and spoken word planning).

In addition to these reaction time effects, the cognitive demands of single word planning have also been demonstrated through variation in pupil dilation. In general, pupil dilation has been demonstrated to be a reliable indicator of cognitive load, as demonstrated, for example, by the increased dilation for every additional item held in memory (e.g. Kahneman & Beatty, 1966). In the linguistic domain its capability to measure cognitive load has been assessed by Papesh and Goldinger (2012). Those authors used a delayed naming paradigm with written words, varying in frequency. However, they cued participants to say "blah" instead of the real written words on some portion of the trials, thereby matching articulatory processes. Despite this control, they observed that planning low frequency words invoked larger pupil dilation than planning high frequency words. Retrieving low frequency words is thus more cognitively demanding than retrieving high frequency words. In addition, it suggests that pupil dilation may be a sensitive measure of cognitive load in language-related tasks.

In sum, both speech perception and speech planning rely, to some extent, on central processing resources. There are two alternatives for how this common dependency may affect both processes in an overlap situation: a) speech planning and speech perception will have to compete for central processing resources and this competition for resources will negatively impact one or both processes; b) speech planning and speech perception will not compete for central processing resources because sufficient resources are available to be shared between both processing streams. This may be because neither process places great demands on attention, and/or because participants are highly skilled at producing and perceiving speech in tandem.

Paradigm and dependent measures.

Two experiments investigated how well listeners encoded what they heard while planning speech. In the exposure phase participants heard a word and at the same time either (1) saw a picture which they had to name (plan task), or (2) saw a meaningless line drawing (which was always the same) which they did not have to respond to (no-plan task). In Experiment 1, participants were instructed to ignore the auditory words. In Experiment 2 they were instructed to also pay attention to the auditory words since they would later participate in a recognition memory task regarding these auditory items. Comparing these differing task goals would allow for an assessment of the impact of intentionally sharing (Experiment 2) or not sharing (Experiment 1) resources between speech planning and perception.

To evaluate the effect of speech planning on the processing of the heard words, we determined how well the words were remembered in an off-line recognition memory task. This recognition memory performance was analyzed with a sensitivity measure of signal detection theory (the sensitivity signal detection index d'[1], see Green & Swets, 1966; Macmillan & Creelman, 2005). In addition, we measured decision speed because correct responses (Hits and Correct Rejections, in detection theory terms) have been shown to result in faster decisions than incorrect responses (Misses and False Alarms: Montefinese, Ambrosini, Fairfield, & Mammarella, 2013; see also Inaba, Nomura, & Ohira, 2005).

Moreover, words that have been attended to during exposure are also recognized faster than words that have received less attention during exposure (Bentin, Kutas, & Hillyard, 1995) and, in some cases, the probability of retrieving an item in a memory test has been shown to remain unaffected by a concurrent task, while speed of retrieval was still affected (Baddeley, Lewis, Eldridge,and Thomson, 1984). Hence, in addition to overall memory accuracy (i.e., d'), decision speed was considered to signal how cognitively demanding the recognition of an item in each task (plan task, no-plan task) was.

Pupillometry in the exposure phase. In addition to behavioral measures, Task-Evoked Pupillary Responses (TEPR) were measured to obtain fine-grained information regarding the resource demands imposed by speech planning as well as those involved in memory encoding and retrieval. TEPRs are phasic changes in pupil size that occur as a consequence of task demands. They are considered psychophysiological markers for task-evoked cognitive activation (e.g., Beatty & Lucero-Wagoner, 2000; Kahneman, 1973; Laeng, Sirois, & Gredeback, 2012). To exemplify, simultaneous interpreting of difficult words has been shown to result in larger pupil dilation than that of easy words (Hyönä, Tommola, & Alaja, 1995) and listening to degraded spoken sentences also modulates pupil dilation (Zekveld, Heslenfeld, Johnsrude, Versfeld, & Kramer, 2014). Similarly, Kuchinke, Võ, Hofmann and Jacobs (2007) demonstrated that visually presented low frequency words led to larger pupil dilation than high frequency words in a lexical decision task. Thus pupil dilation is a suitable measure for monitoring cognitive demand in our study.

Pupillometry in the recognition memory test. Recent literature has also linked pupillary dilation patterns during encoding and retrieval to recognition memory performance. A key finding has been the pupil old/new effect, which is the observation that old items in a recognition memory task induce larger pupil dilation than new items. Even though the "classic" pupil old/new effect involves comparing hits to correct rejections (CR; Võ et al.,

2008) the notion has been extended to other related situations (e.g. Goldinger & Papesh, 2012; Kafkas & Montaldi, 2012; Papesh, Goldinger, & Hout, 2012). In particular, this pupillary effect has not only been found during retrieval in the recognition memory task, but also during encoding of the stimuli, with items that elicited an "old" response in the recognition memory test also eliciting a larger dilation in the exposure phase (Papesh et al., 2012; but see also Kafkas & Montaldi, 2011). Moreover, deep processing instructions have been shown to result in higher accuracy and larger pupil dilation during retrieval in the recognition memory test, compared to shallow processing instructions (Otero, Weekes, & Hutton, 2011). Thus in the recognition memory task we use pupil dilation as an additional index of recognition memory strength.

Predictions

Exposure phase. The cognitive effort put into planning the names of pictures should be visible in the pupillary response, with the plan task leading to larger dilation than the no-plan task. To control whether participants complied with the fast naming instructions, we included a picture frequency manipulation. If participants tried to name pictures fast, high frequency pictures should be named faster (e.g. Janssen, Schirm, Mahon, & Caramazza, 2008; Jescheniak & Levelt, 1994; but see Goldinger, Azuma, Abramson, & Jain, 1997; Papesh & Goldinger, 2012), and perhaps more accurately, than low frequency pictures. Motivated by the findings of Hyönä and colleagues (1995) and Kuchinke and colleagues (2007) we also expected larger pupil dilation when participants named low frequency pictures than when they named high frequency ones. For the no-plan task[2] we expected a pattern similar to Papesh and colleagues, (2012) according to which items correctly recognized in the recognition memory test would be accompanied by larger pupil dilation during the exposure phase than items that would later be missed.

Recognition memory test. The recognition memory test provided the measures to evaluate the quality of encoding during exposure. Prior research has shown that recognition memory is poorer when attention had been divided during encoding (see for example, Fernandes and Moscovitch, 2000). We predicted that auditory targets encoded during planning a single word would lead to inferior recognition memory accuracy compared to items encoded during no speech planning.

We further anticipated a general modulation of pupil size by memory strength. That is, "old" responses to auditory targets (hits) should lead to overall larger dilation than "new" responses (misses). In addition, we expected modulation of this effect by the encoding depth during the exposure phase. Auditory targets that were encoded during speech planning, which is more likely to result in shallow encoding (see in Otero, Weekes, & Hutton, 2011), should result in smaller dilation than auditory targets that were encoded during no planning (deep encoding).

The consequences of having to distribute limited resources and thereby achieving poorer encoding would also be evident in decision speed, with auditory targets of the no-plan task leading to faster response than plan auditory targets. In accordance with Montefinese and colleagues (2013) - and following the assumption that decision speed reflects the consequences of how cognitively demanding encoding was - we also expected this effect to be modulated by response type and be more prominent for "old" (hits) than for "new" (misses) responses to auditory targets. "Old" responses to no-plan stimuli (no-plan hits) would be given faster than "old" responses to plan stimuli (plan hits), while we should observe no such difference for "new" responses (plan and no-plan misses). In contrasting decision speed between lure items (correct rejections) and plan and no-plan targets we expected that no-plan hits would be decided upon faster than lures, whereas plan hits would be decided upon with a speed more similar to that for lures. Thus, plan targets would be treated more like lures compared to no-plan targets.

To summarize, in two experiments we asked first whether sharing resources between speech planning and listening had consequences for memory of the heard words, and we assessed whether any observed effect of this setup is moderated by promoting intentional (Experiment 2) rather than incidental (Experiment 1) encoding of the heard words. Primary dependent measures of encoding during perception are evaluated in conjunction with secondary measures on production performance, to provide a complete profile of resource sharing during this dual processing task.

Experiment 1

Method

Overview. The experiment consisted of an exposure phase and a recognition memory test. In the exposure phase, the participants heard spoken words accompanied either by pictures they had to name (plan task) or by a meaningless line drawing (no-plan task). The pictures were either low or high in name frequency.

The recognition memory test was a surprise test for the words heard during the exposure phase presented amidst an equal number of lures. We compared recognition memory accuracy, decision speed, and pupil dilation for words that had been presented on plan versus no-plan trials of the exposure phase. Decision speed for targets in the plan and no-plan task was also contrasted to that for lures. This comparison was informative on how similar or different the targets of the two tasks (plan, no-plan) were treated compared to lures. We also assessed pupil dilation in the exposure phase for no-plan items subsequently remembered or forgotten in the recognition memory test.

To create a delay between the exposure phase and the recognition memory test participants were asked to carry out a filler task. This was the “number-letter” task, which is designed to capture the ability to shift between tasks or mental sets (Miyake et al., 2000; Rogers & Monsell, 1995). We examined whether performance on this task correlated with response speed in the exposure phase or with performance in the recognition memory test of the experiment. As this turned out not to be the case, the procedure in the number-letter task and the results are not further described.

Both experiments also included a naming phase that was carried out after the recognition memory test. Here participants named the same pictures as in the exposure phase intermixed with an equal number of new pictures. The results are not reported because they are irrelevant for the specific research question discussed in the current chapter.

Participants. The experiment was carried out with 32 paid participants from the participant pool of the Max Planck Institute for Psycholinguistics, Nijmegen. They were adult native speakers of Dutch (mean age = 22.7 years, 24 female). Most of them were university students. All participants had normal or corrected-to-normal vision. No participant reported any speech or hearing problems or being diagnosed with dyslexia. One participant had to be excluded from the analysis for not conforming to the exposure phase instructions and another because of technical errors. Consent for conducting the study had been obtained from the Ethics Board of the Social Sciences Faculty of the Radboud University Nijmegen.

Materials and Design. In the exposure phase, all participants heard the same 104 spoken words. Four words were used on practice trials. The remaining words were randomly assigned to one of two lists (list 1 and list 2) of 50 words each. For half of the participants (group 1) the words of list 1 were accompanied by a meaningless line drawing (see Figure 1), while the words of list 2 were accompanied by pictures they had to name. For the remaining participants (group 2) the spoken words of list 1 were combined with pictures and those of list 2 with the meaningless line drawing (see Figure 2 for a visualization of the counterbalancing). As a result, each auditory target appeared in both tasks, thereby controlling for any effect of differences in the auditory targets.

Figure 1. The meaningless line drawing used in the no-plan task of the exposure phase (Experiments 1 and 2).

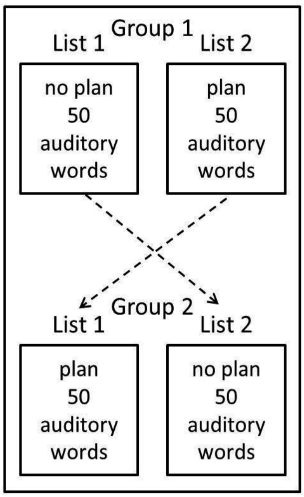

Figure 2. Visualization of the counterbalancing of the two auditory word lists across tasks and participant group.

The words were monosyllabic nouns selected from the CELEX database (Baayen, Piepenbrock, van Rijn, 1993). The words in the two lists (plan and no-plan items) were matched for spoken duration, lemma frequency and number of phonemes (see Table 1 for averages and Appendix B for a full listing).

Table 1

Characteristics (means with standard deviations in parentheses) of auditory words in Lists 1 and 2 of the exposure phase and of lures (recognition memory test). "Overall Targets" lists the mean values of Lists 1 and 2 for comparison to the "Lures".

characteristic	List 1	List 2	Overall Targets	Lures
audio duration (ms)	646 (90)	670 (110)	658 (101)	663 (96)
log frequency	3.68 (0.89)	3.89 (0.88)	3.78 (0.88)	3.75 (0.83)
number of phonemes	3.44 (0.54)	3.46 (0.50)	3.45 (0.52)	3.45 (0.52)

The pictures were 102 line-drawings selected from the data base provided by Severens, Van Lommel, Ratinckx, and Hartsuiker, 2005. Two pictures were used on practice trials, and 50 each on the plan trials of participant group 1 and 2, respectively. Thus, the two groups of participants saw different sets of pictures. This was the case because, as already mentioned in the overview section of Experiment 1, the study included a final naming phase, in which participants named a mixture of pictures they had named in the exposure phase with new pictures. Old and new pictures were counterbalanced within the two groups of participants. Pictures named in the exposure phase by group 1 were named in the final naming phase by group 2 and vice versa.

Half of the pictures in each set had high frequency names and half had low frequency names (List 1, high- vs low-frequency: $t(35.989) = 12.0548$, $p < .001$; List 2, high- vs low-frequency: $t(34.019) = 9.8627$, $p < .001$). The four sets of pictures (high and low frequency pictures presented to groups 1 and 2) were matched in terms of average name frequency, name agreement (*H*-statistic), age of acquisition, length in number of syllables and number of phonemes, visual complexity (as specified by Bates et al, 2003) and luminosity (luminosity was measured using the luminosity tool in Adobe® Photoshop®, Version 11.0.2). Table 2 lists the characteristics of the group 1 and group 2 picture sets with the corresponding information for high and low frequency pictures per group (see Appendix C for a listing of the pictures). The pictures were combined with spoken words into semantically and phonologically unrelated pairs.

In the recognition memory test, all participants heard the same stimuli, namely the 100 words used in the exposure phase (targets hereafter), and an additional 100 lures, which were also monosyllabic nouns drawn from the CELEX data base. Two more lures were added to be used in the practice of the recognition memory test, which also included two of the practice items from the exposure phase. Lures and targets were matched for spoken duration, lemma frequency and number of phonemes (see Table 1 for averages and Appendix D for a listing of the lures). In creating the exposure phase and recognition memory test audio materials we had initially selected 200 nouns that matched our criteria and then randomly assigned those to a target or lure role.

Table 2

Characteristics (means with standard deviations in parentheses) of high and low frequency picture names in Group 1 and Group 2

characteristic	Group 1 Pictures			Group 2 Pictures		
	Frequency		overall	frequency		overall
	high	Low		High	low	
Age of Acquisition (AoA)	5.54 (0.91)	7.33 (1.11)	6.28 (1.33)	6.27 (1.21)	7.75 (1.87)	6.79 (1.62)
log frequency	0.81 (0.27)	0.084 (0.14)	0.45 (0.42)	0.75 (0.32)	0.07 (0.13)	0.42 (0.42)
luminosity	242.44 (6.98)	238.99 (8.33)	240.75 (7.79)	234.15 (14.65)	239.33 (7.62)	236.58 (12.04)
name agreement (H-statistic)	0.76 (0.75)	0.94 (0.51)	0.85 (0.64)	0.70 (0.53)	0.95 (0.59)	0.82 (0.57)
number of phonemes	6.48 (1.58)	7.48 (2.29)	6.98 (2.01)	6.88 (1.84)	7.08 (1.72)	6.98 (1.77)
number of syllables	2.44 (0.71)	2.72 (0.74)	2.58 (0.73)	2.42 (0.76)	2.62 (0.77)	2.52 (0.76)
visual complexity	13656 (5967)	16068 (68890)	14838 (6483)	20121.12 (12081)	14860 (4924)	17652 (9704)

Note. The luminosity value listed here refers to the to be named pictures. The line drawing used for the no-plan task had a luminosity value of 238.68.

Each spoken word was presented together with a six-point scale showing the numerals 1 to 6, with the value 1 labeled "nieuw" (new) and the value 6 labeled "oud" (old).

The order of the targets was the same as in the recognition memory test, but 0 to 4 lures intervened between successive targets.

The auditory target words and lures were recorded in a quiet booth by a male native speaker of Dutch. He read the words from a list featuring targets and lures in random order, without knowing which category each word belonged to. The recorded list was later spliced into individual sound files for each word.

The pictures were presented as black line drawings on a white background, scaled to fit into frames of 300 by 300 pixels (4.5° of visual angle from the participant's position).

Apparatus. The experiment took place in a dimly lit sound-treated booth. Eye position and pupil dilation were measured in a monocular tower mount setup using an SR-Research Eyelink 1000 system, with a sampling rate of 500 Hz. The participant was seated at a distance of 1 meter in front of an Acer flat screen (resolution of 1024 x 768 pixels) with the head resting on chin and forehead rests. Stimulus presentation was controlled by SR Research Experiment Builder software, version 1.6.121. The spoken words were presented using Sennheiser HD201 Lightweight Over-Ear Binaural headphones.

Procedure. At the beginning of the session chin and forehead rests of the eye-tracker were adjusted. The participants received written instructions on screen about the exposure phase. They were asked to name all pictures correctly and as quickly as possible while ignoring the spoken words. After the participant confirmed having understood the task, the eye tracker was calibrated using a nine-point calibration procedure.

Every trial started with the presentation of a fixation cross for 3000 ms followed by synchronous onset of a word and line drawing, which was either a picture or the meaningless drawing. In order to induce fast speech planning, the visual stimulus was presented for only 250 ms and was followed by a blank interval of 1750 ms. The long exposure to the fixation cross and a blank interval were necessary to allow for pupil size to return to baseline between trials. Reaction Times (RTs) were measured from picture onset on (note that picture and auditory word had a synchronous onset).

The experiment began with four practice trials, two of which featured pictures that had to be named. After the practice trials the participant was given the opportunity to ask any questions. Then the experimental trials began. There were four blocks of 25 trials each, separated by short pauses. At the beginning of each block a drift correction was performed and if necessary the eye tracker was recalibrated. The exposure phase took approximately 12 minutes to complete. Then the participants completed the number-letter task mentioned above. This took approximately 20 minutes.

During the following recognition memory test participants listened to the sequence of targets and lures described above and indicated for each of them how certain they were that the item had, or had not, been heard in the exposure phase. At the beginning of each trial, a fixation cross was shown for 2000 ms, followed by simultaneous onset of a word and the rating scale. The scale remained on screen for 3000 ms and was followed by a blank interval of one second. Participants had to look at the scale and issue a confidence estimate by saying out loud one of the numbers ranging from 1 (indicating certainty that the item was new) to 6 (indicating certainty that the item was old). They were encouraged to use intermediate numbers to indicate uncertainty. The decision latencies were recorded by voicekey. The auditory word onset served as the timepoint from which on decision latencies were measured. There were four practice trials followed by eight test blocks of 25 experimental items each. There were short pauses between blocks. The recognition memory test took approximately 25 minutes to complete.

Recording and analysis of pupil size. Pupil size was recorded throughout the experiment. In Eyelink 1000 pupil size is reported in arbitrary units with a linear relation to the recorded pupil diameter (see Eyelink User Manual and Einhäuser, Stout, Koch, & Carter, 2008). The acquired signal was processed via in-house code developed in R mainly by the first author (version 2.14.2; The R foundation for statistical computing).

First the signal was down-sampled to 50 Hz. Next the change in pupil size from sample-to-sample was evaluated in order to identify and eliminate outliers, typically due to blinks. To this end, any samples within a trial with a value deviating more than one standard error from the mean sample-to-sample pupil size change of that trial were considered outliers and were replaced by a missing value (see Piquado, Isaacowitz, & Wingfield, 2010 and Skaftnes, 2012 for similar procedures). Trials where more than 25% of the samples were missing were excluded from further analyses. Missing samples of the remaining trials were replaced via linear interpolation. Moreover, for every phase, any participants who lost more than 25% of their trials in one of the experimental conditions (including data loss due to erroneous or late responses) were excluded from further analyses. This was the case for five participants in the exposure phase and two in the recognition memory test.

To correct for tonic changes in pupil size and for differences in average pupil size across participants, the absolute pupil dilation was converted into relative pupil dilation following van Rijn, Dalenberg, Borst, & Sprenger (2012; see also Montefinese et al., 2013). This was done by baseline-correcting and normalizing pupil size on a trial-by-trial basis. The baseline was the average pupil size recorded during the last second of the presentation of the fixation cross in a trial. The critical time window began with the onset of the stimulus (i.e., the picture or meaningless line drawing and the spoken word in the exposure phase and the rating scale and the spoken word in the recognition memory test) and had a duration of 2000 ms in the exposure phase and of 3000 ms in the recognition memory test. To compute the relative pupil dilation, the baseline pupil size of a trial was subtracted from the recorded pupil size for each sample of that trial. The resulting difference score was divided by the baseline size and multiplied by 100. The result is the percentage of pupil diameter change (PDC) due to the experimental manipulation and is used to plot the TEPR. Finally, the mean pupil size per trial was computed as the average of the PDC values across the critical time window. We also report the peak pupil size, which is the maximum PDC in the critical time window. According to Beatty and Lucero-Wagoner (2000) peak pupil size is more susceptible to random variations in the signal than mean pupil size, but more independent of number of observations in the critical time-window.

Results

Exposure phase

Behavioral Data. To evaluate the influence of picture LogF on error rates in the plan task of the exposure phase we removed responses equal to or faster than 300 ms from the analyses (2.6 % of the data). Error rates were evaluated using generalized linear mixed-effects regression models in R (version 2.14.2; The R foundation for statistical computing; lme4 package; Bates et al., 2015) and were fitted with binomial distributions (Jaeger, 2008). We followed a design-driven approach and determined the random structure of the models following the suggestions of Barr, Levy, Scheepers, and Tily (2013). Interactions were evaluated using a backward elimination procedure that established whether including a specific interaction in the model improved model fit in describing the data. In each step non-significant interactions were removed from the equation. All additive predictors justified by the design were kept in the models. In addition, a null model comprising only the random structure of the models was established. The control factor experimental list was also included in all analyses but removed when the analysis indicated that it did not add significantly to the explanatory power of the model. All models, including the null model were compared using log-likelihood ratio tests to assess model improvement. The optimal (best fitting) models are reported.

More naming errors occurred for low frequency pictures (19.04 %) than for high frequency pictures (5.44 %). The full model for error rates included picture LogF (high and low; fitted as -1 and 1, respectively) and list (List 1 and List 2, fitted as -1 and 1, respectively) as fixed factors, random intercepts for participants and picture-names, and a random slope for picture LogF by participants[3]. Picture LogF contributed significantly to the error rates in the exposure phase ($b = .937$, $p < .001$, intercept = -3.086).

To evaluate response speed in the exposure phase as a function of picture LogF we removed errors (12.3% of the data). Outliers were set to two standard deviations from the participant mean resulting in elimination of a further 5% of the data.

High frequency pictures were named faster (mean = 1032 ms) than low frequency pictures (mean = 1101 ms). The random structure of the models of response speed included a random intercept and random slope for picture LogF by participants and a random intercept for picture-names. The full model specification included the fixed factors picture LogF (high and low, fitted as -1 and 1 respectively) and list (List 1 and List 2, fitted as -1 and 1 respectively). The model that best fit the response speed data only included a significant main effect of picture LogF ($b = 35.66$, $p = .005$, intercept = 1079.24).

Pupil Size Data. We evaluated the mean and peak pupil size in the two tasks to derive an estimate of the cognitive demands invoked by planning while listening. Planning speech while listening to speech induced larger pupil dilation ($Plan_{mean} = .0194$[4], $Plan_{peak} = .1186$) than listening ($No\text{-}Plan_{mean} = .0069$, $No\text{-}Plan_{peak} = .0525$). The optimal model included a fixed effect of task (plan and no-plan, fitted as -1 and 1 respectively; $b_{mean} = -0.0204$, $p = .004$, intercept =.0248; $b_{peak} = -0.0339$, $p = .009$, intercept =.0863) and a random intercept by participants and picture-names, with a random slope for task by participants.

We then evaluated how the cognitive demand invoked by planning while listening was modulated by ease or difficulty of planning as reflected in picture log frequency. Peak pupil dilation was entered as the dependent variable in a model that included picture LogF (high and low, fitted as -1 and 1 respectively) and list (List 1 and List 2, fitted as -1 and 1 respectively) as fixed factors, random intercepts for participants and picture-names and a random slope for picture LogF. The model that best fit the peak pupil size data included picture LogF (b =.0083, p =.011, intercept =.1204) as a fixed factor. Thus low frequency picture-names induced larger pupil dilation than high frequency picture-names. The same model was also tested using mean pupil size as the dependent variable (rather than peak size) but did not reveal any significant effect of LogF (b =. 0020, p =.338, intercept =.0454).

Recognition memory test

Behavioral Data. The main focus of the recognition memory test was to assess whether participants would differ in their recognition memory accuracy as a function of encoding task (plan, no-plan). To quantify recognition memory accuracy we computed d', which is the signal detection index for sensitivity (Macmillan & Creelman, 2005). By asking participants to provide graded (6-point scale) estimates we obtained empirical ROC curves for the sensitivity data. The group-level ROC curves are depicted in Figure 3. This is a way to consider the possibility that the participant could have maintained different response criteria during the task (see also Papesh et al., 2012). Sensitivity can then be estimated separately for each criterion. To do so we split the 6-point scale into 5 split-categories of two groups each, with a notional criterion line shifting from number to number, left to right (see the legend of Figure *3* for details on all five resulting splits). If a participant for example maintained a conservative criterion at all times, then she would use only 6 to indicate "old" and 12345 to indicate "new". Of all the resulting split-categories the most intuitive would of course be the one implied by the instructions the participants received; namely to group numbers 1, 2, and 3 as a "new" and 4, 5 and 6 as an "old" response, thus 123_456.

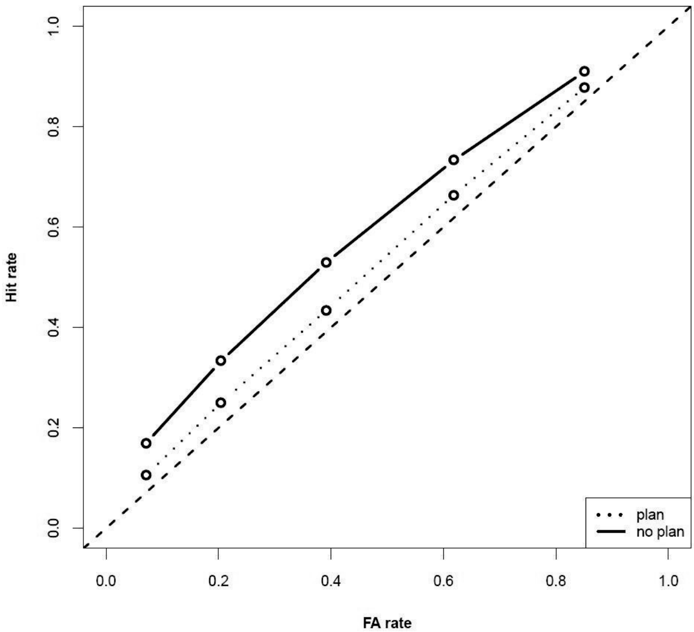

Figure 3. Group ROC curves for Experiment 1. The x-axis represents False Alarm (FA) rate, and the y-axis Hit (H) rate. The 5 circles on each task line correspond to the 5 different scale-splits (from left to right: 12345_6, 1234_56, 123_456, 12_3456, 1_23456. The numbers that appear left of the underscore represent "new"- and the numbers right of the underscore represent "old"- responses). The dashed diagonal line represents chance level. Values in the upper right corner of the graph represent a liberal criterion whereas values in the lower left corner of the graph represent more conservative criteria.

We used the d' scores as estimated by the 123_456 - split to conduct a paired-samples t-test. On average, participants' recognition performance was higher in the no-plan ($M_{\text{No-plan}}$

=.367, SE $_{No\text{-}plan}$ =.037) than in the plan task (M $_{plan}$ =.109, SE $_{plan}$ =.037; t (29) = 5.712, p <.001, r =.727). Even though these scores might be considered rather low d' values, one should keep in mind that participants had to discriminate between 200 monosyllabic auditory items (100 targets and 100 lures), and received no warning at exposure about the subsequent recognition memory task. In light of that it is rather striking that their performance was better than chance level not only in the no-plan task (for all splits: 1_23456: t_1 (21) = 6.085, *p <.001*; 12_3456: t_1 (29) = 7.542, *p <.001*;123_456: t_1 (29) = 9.904, *p <.001*; 1234_56: t_1 (29) = 7.884, *p <.001*; 12345_6: t_1 (22) = 7.894, *p <.001*;) but also in all but one split-category of the plan task (1_23456: t_1 (22) = 1.383, *p* =.181; 12_3456: t_1 (29) = 3.118, *p* =.004; 123_456: t_1 (29) = 2.993, *p* =.006; 1234_56: t_1 (29) = 3.24, *p* =.003; 12345_6: t_1 (20) = 3.912, *p* =.001).

Next we evaluated decision speed as a function of task (plan, no-plan) and recognition memory value. We removed responses equal to or faster than 300 ms from the analyses (7% of the data). Outliers were set at two standard deviations from the participant mean resulting in elimination of a further 4% of the data. Decision speed was entered as the dependent variable in a model that included task (plan and no-plan, fitted as 1 and -1 respectively), recognition memory value (6-point scale centered), list (List 1 and List 2, fitted as -1 and 1 respectively), and the interaction of task with recognition memory value as fixed factors. The random effects structure included random intercepts for participants and spoken words as well as random slopes (by participants and by spoken words) for task and for recognition memory value and the corresponding interactions.

The model that best fit the response speed data included significant main effects of task (b = 20.260, p =.022, intercept = 1906.836) with no-plan items leading to faster decision than plan items (No-Plan $_{mean}$ =1871 ms, No-Plan $_{standard\ deviation}$ = 236; Plan $_{mean}$ =1921 ms, Plan $_{standard\ deviation}$ = 232), recognition memory value (b = -19.293, p =.025) with high recognition memory values leading to faster decision than low recognition memory values, as well as a significant interaction of task by recognition memory value (b = 16.926, p =.003). The interaction was evaluated by separate mixed effects analysis of "new" and "old" recognition memory values (123_456 split) with task as a fixed factor and random intercepts for participants and spoken words as well as random slopes for task for participants and spoken words. It revealed that "new" responses (misses) were given equally fast for plan and no-plan items (b = 5.527, p =.626, intercept = 1917.219; Plan-New $_{mean}$ =1924 ms, Plan-New $_{standard\ deviation}$ = 261; No-Plan-New $_{mean}$ =1913 ms, No-Plan-New $_{standard\ deviation}$ = 288) whereas an "old" response to no-plan items (no-plan hit) was given faster than an “old” response to plan items (plan hit; b = 42.67, p =.001, intercept = 1891.25; Plan-Old $_{mean}$ =1929 ms, Plan-Old $_{standard\ deviation}$ = 223; No-Plan-Old $_{mean}$ =1854 ms, No-Plan-Old $_{standard\ deviation}$ = 210).

An additional analysis evaluated how decision speed for lures related to decision speed for targets in the two tasks. This comparison showed how the targets of the two tasks (plan, no-plan) were responded to in comparison to lures. Task (lure, plan, no-plan), response (new, old; following the 123_456 split) and list were entered in the model. "Lure-old response" was set as the reference level. Note that the combination of the factors task and response results in the four possible outcomes of Signal Detection Theory: plan and no-plan old are "hits", plan and no-plan new are "misses", lure old is a "false alarm" and lure new is a "correct rejection". The models had task, response, and the interaction of task by response as fixed factors and a random intercept by participants and spoken words as well as random slopes for task and response. The model that best fit the data indicated that for “old” responses, no-plan items were decided on faster than lures (b $_{no\text{-}plan\ old}$ = -100.752, p <.001, intercept = 1948.210; No-Plan-Old $_{mean}$ =1854 ms, No-Plan-Old $_{standard\ deviation}$ = 210; Lures-Old $_{mean}$ =1949 ms, Lures-Old $_{standard\ deviation}$ = 266), whereas plan items did not differ from lures (b $_{plan\ old}$ = -15.639, p =.489; Plan-Old $_{mean}$ =1929 ms, Plan-Old $_{standard\ deviation}$ = 223). A significant interaction of response with the no-plan task (b = 78.676, p =.001) was due to the fact that the difference between no-plan items and lures was bigger when labeling the item as "old" (-101 ms) compared to "new" (-22 ms). The interaction of response with the plan task was not significant ($b = 6.139$, $p = .802$) since the difference between plan items and lures was not bigger for "old" (-16 ms) as compared to "new" responses (-9 ms). Thus plan items were decided upon in a similar way to lures, whereas no-plan items and, especially the ones that triggered an "old" response (hits), were decided upon faster.

Pupil Size Data. "Old" responses were accompanied by larger dilation than "new" responses. Two mixed effects models -for mean and peak pupil size respectively- were assessed. Both initial models included fixed effects for task (plan and no-plan fitted as -1 and 1 respectively) and recognition memory value (6-point scale centered) as a simple numeric term, but also its quadratic part, in an attempt to model the u-shaped curve of the pupil size data (Figure 4). The interaction term of task by response was also part of the fixed structure. The random structure of the models included random intercepts by participants and spoken words, as well as random slopes for task and response and their interaction by participants and spoken words. The final models included fixed effects of the simple (mean pupil size: $b = .0023$, $p = .017$, intercept $= .0214$; peak pupil size: $b = .0047$, $p = .001$, intercept $= .0700$) and the quadratic term of response (mean pupil size: $b = .0024$, $p < .001$; peak pupil size: $b = .0046$, $p < .001$) and non-significant fixed effects of task (mean pupil size: $b = .0002$, $p = .886$; peak pupil size: $b = .0003$, $p = .893$), with random intercepts for participants and spoken words, along with random slopes for task and response by participants and spoken words. Thus, only response but not task (plan, no-plan) affected pupil dilations.

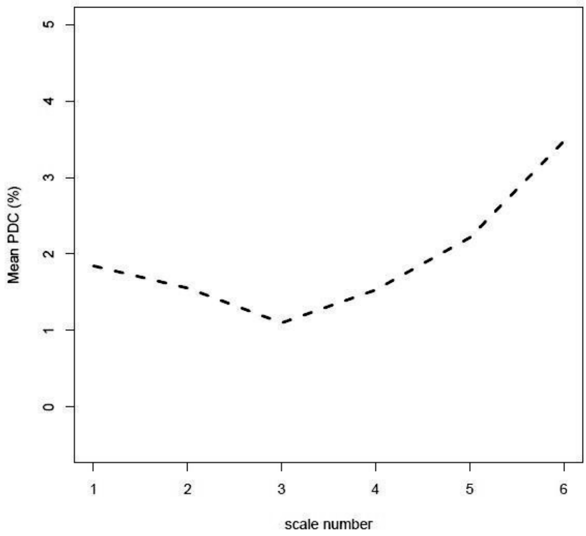

Figure 4. Average percentage of pupil diameter change (PDC) as a function of scale number (1 = "new", 6 = "old) in the recognition memory test for experiment 1. The x-axis represents response number (scale). Note that the pupil size data display a u-shaped curve that prompted us to include also a quadratic term in our model.

Following Papesh and colleagues (2012) we evaluated whether the encoding pattern for the spoken words presented in the no-plan task of the exposure phase – as measured via pupil dilation - would predict later correct recognition memory value (6-point scale) for these items in the recognition memory test. Plan trials were not evaluated in this analysis. The two initial models included mean or peak pupil size as a fixed factor and random intercepts by participants and spoken words. Correct recognition of a word in the recognition memory test was not linked to larger dilation during encoding in the exposure phase, as was evident by both mean ($b = 1.497$, $p = .083$, intercept =.1932) and peak ($b = 1.029$, $p = .175$, intercept =.1459) dilation estimates. Note however that there was a trend towards significance for mean pupil size.

Discussion

In Experiment 1 we sought to evaluate how well listeners encoded what they heard while planning speech. Encoding quality was assessed in a recognition memory test using the d' sensitivity measure of signal detection theory, decision speed and TEPRs. The resulting sensitivity indices revealed that in both tasks (plan, no-plan) performance was better than chance level. In addition they supported our hypothesis that auditory targets encoded during planning a single word would lead to inferior recognition memory accuracy than items encoded during no speech planning.

In the pupil dilation data recorded during the recognition memory test, we observed the pupil old/new effect[5] (Papesh et al., 2012) with "old" responses to auditory targets (hits) leading to larger dilation than "new" responses (misses), illustrating that pupil dilation reflects memory processes (be it retrieval per se or confidence). However, pupil size at test was not further modulated by encoding depth during the exposure phase. That is, auditory targets encoded during no speech planning (deep encoding) did not result in larger dilation at test than auditory targets encoded during speech planning (shallow encoding). Our manipulation might not have been sensitive enough to pick up this effect. We return to this issue in Experiment 2.

The effects of encoding depth were however visible in the decision speed. As expected, no-plan targets led to faster decision than plan targets, and high recognition memory values (6-point scale), which were linked to "old" responses, led to faster decision than low recognition memory values (linked to "new" responses). The interaction of task with recognition memory value revealed that it was the higher recognition memory values ("old" responses) that were given faster in the no-plan task compared to the plan task, whereas lower memory values ("new" responses) were given equally fast for plan and no-plan items.

This suggests that accessing the memory trace of plan items was more demanding than accessing the memory trace of the no-plan task.

In fact, with respect to RTs, plan items seem to have been treated similar to lures. In comparing targets to lures we found that no-plan hits were decided upon faster than false alarms ("old" response to a lure), while this was not the case for plan hits, which were decided upon in similar speed as false alarms. Moreover responding that an item is "old" resulted in a larger difference in decision speed between lures (false alarms) and no-plan items (no-plan hits) compared to responding that the lure or no-plan item is "new" (correct rejection and plan miss respectively). This was not the case when comparing speed of "old" to "new" response decisions between lures and plan items.

Response speed and accuracy in the exposure phase was modulated by the log frequency of the picture-name, with high frequency picture-names leading to faster responses and higher accuracy than low frequency ones. This confirmed that participants were indeed adhering to our instruction of fast planning.

Pupil dilation measures acquired during the exposure phase established a continuous measure of the cognitive demands of planning speech (while listening to speech), with the plan trials inducing larger dilation than no-plan trials. This cognitive demand of planning seems to have drawn away resources from listening. The fact that pupil dilation in the exposure phase reflects the processing load of planning is also supported by the finding that low frequency picture-names induced a larger dilation than high frequency picture-names. At the same time pupil dilation patterns of the no-plan trials during this phase were not predictive of later recognition memory value in the recognition memory test (even though there was a tendency towards significance).

Experiment 2

In Experiment 1 we had instructed participants to focus on picture naming and ignore the auditory words, thereby relying on their incidental encoding capacity. It could well be that our finding whereby plan auditory targets are remembered less well than no-plan auditory targets is mainly driven by our instruction and not by the need to accommodate two processing streams at the same time. To assess whether this was the case, we conducted Experiment 2, which involved only an instruction change relative to Experiment 1. Now participants were explicitly instructed in the exposure phase to name the pictures but also to listen to the auditory words and try to remember them, because they would later be asked to indicate how sure they were that they had heard a given word in the exposure phase.

Method

Participants. 39 participants from the Max Planck Institute for Psycholinguistics participant pool took part in the study (mean age = 21.9 years, 31 female). None had previously participated in Experiment 1. All were native speakers of Dutch and had normal or corrected-to-normal vision. None of them reported a speech or hearing problem, and none had been diagnosed with dyslexia. Participants gave written consent on their participation and received monetary compensation. Two male and four female participants had to be excluded from the analysis, due to either response bias in the recognition memory test (participants with lambda values below 0 or above 1.5 were excluded; Wickens, 2001), not conforming to the recognition memory test task instructions and prerequisites or due to technical issues.

Materials. Materials were identical to Experiment 1.

Apparatus. The apparatus was identical to Experiment 1.

Procedure. The procedure was identical to Experiment 1. Only one change was introduced in the exposure phase. Participants were instructed to also attend to the auditory stimuli because they would later be asked to indicate how confident they are that they have heard some words in the exposure phase.

Recording and analysis of pupil size. The analysis protocol was identical to Experiment 1.

Results

Exposure phase

Behavioral Data. Responses equal to or faster than 300 ms were removed from the analyses (3.2% of the data). As can be seen on the right panel of Figure 8, error rates were higher for low frequency pictures (19.80 %) than for high frequency pictures (9.28 %). The full model included picture LogF (high and low fitted as -1 and 1 respectively) and list (List 1 and List 2 fitted as -1 and 1 respectively) as fixed factors, random intercepts for participants and picture-names, and a random slope for picture LogF by participants. Picture LogF (b =.599, p =.001, intercept = -2.599) contributed significantly to the error rates in the exposure phase.

Prior to the analyses of response speed we removed errors (14.6% of the data) and outliers (5.7% of the data). The random structure of the response speed models included a random intercept and random slope for picture LogF by participants and a random intercept for picture-names. The full model specification included the fixed factors picture LogF (high and low fitted as -1 and 1 respectively) and list (List 1 and List 2 fitted as -1 and 1 respectively). The model that best fit the naming latency data included significant main effects of picture LogF (b = 35.18, p =.007, intercept = 1055.62) only. The effect of picture

LogF was due to the fact that high frequency pictures were named faster (mean = 1002 ms) than low frequency pictures (mean = 1073 ms).

Pupil Size Data. Figure 5 depicts the average percentage of pupil diameter change (PDC) as a function of task and, among plan items, picture LogF. Planning speech while listening to speech induced larger pupil dilation ($Plan_{mean}$ =.0168, $Plan_{peak}$ =.1115) than passive listening ($No\text{-}Plan_{mean}$ =.0059, $No\text{-}Plan_{peak}$ =.0535. See also Figure 5). The optimal model included significant fixed effects of task (plan and no-plan fitted as -1 and 1 respectively) for mean pupil size (b = -0.0182, p =.018, intercept =.0214) but only marginally significant for peak pupil size (b = -0.0301, p =.054, intercept =.083) and a random intercept by participants and picture-names, with a random slope for task by participants.

Figure 5 further suggests that pictures with a low LogF resulted in a larger dilation than pictures with a high LogF. To assess this pattern, mean pupil dilation was entered as the dependent variable in a model that included picture LogF (high and low, fitted as -1 and 1 respectively) and list (List 1 and List 2, fitted as -1 and 1 respectively) as fixed factors, random intercepts for participants and picture-names and a random slope for picture LogF. The analysis revealed that LogF (b =.0049, p =.024, intercept =.0402) contributed significantly to mean pupil size. The same model was tested using peak pupil size as the dependent variable. The model that best fit the peak pupil size data included only picture LogF as a fixed factor (b =.0095, p =.009, intercept =.1142).

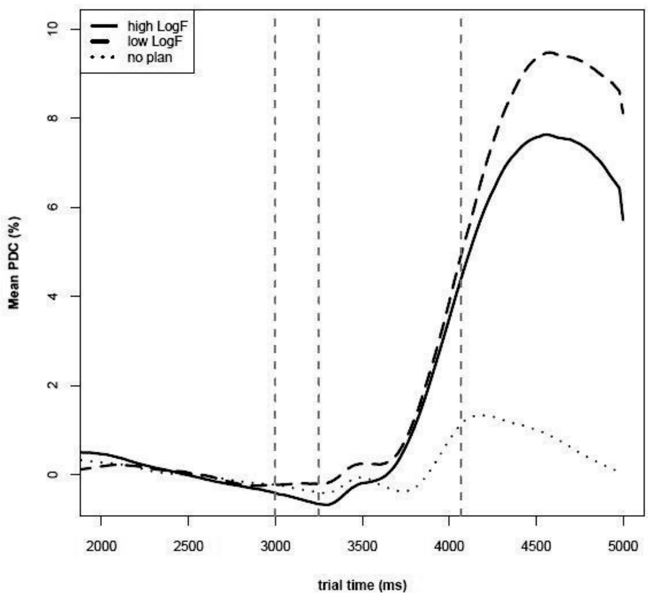

Figure 5. Average percentage of pupil diameter change (PDC) as a function of picture LogF in the plan task of the exposure phase for Experiment 2. The x-axis represents trial time in ms. The vertical dashed lines (starting at the leftmost line) indicate picture and audio onset, picture offset and mean response time. The critical time-window during which pupil size was assessed was from 3000 – 5000 ms. The baseline is from 2000 to 3000 ms. The figure also depicts the PDC data from the no-plan task.

Recognition memory test

Behavioral Data. Figure 6 depicts group-level ROC curves for Experiment 2 in the plan and no-plan task. Paired-samples t-test revealed that, on average, participants' recognition performance as measured with d' was higher in the no-plan ($M_{No\text{-}plan}$ =.498, $SE_{No\text{-}plan}$ =.052) than in the plan task (M_{plan} =.304, SE_{plan} =.050; t (32) = 3.836, p =.001, r =.56). Participants' performance was better than chance level for all splits, both in the no-plan (1_23456: t_1 (25) = 4.019, *p <.001*; 12_3456: t_1 (32) = 9.119, *p <.001*; 123_456: t_1 (32) = 9.654, *p <.001*; 1234_56: t_1 (32) = 9.015, *p <.001*; 12345_6: t_1 (27) = 8.293, *p <.001*) and in the plan task

(1_23456: t_1 (26) = 2.786, *p* =.010; 12_3456: t_1 (32) = 5.323, *p* <*.001*; 123_456: t_1 (32) = 6.050, *p* <*.001*; 1234_56: t_1 (32) = 6.286, *p* <*.001*; 12345_6: t_1 (25) = 5.793, *p* <*.001*).

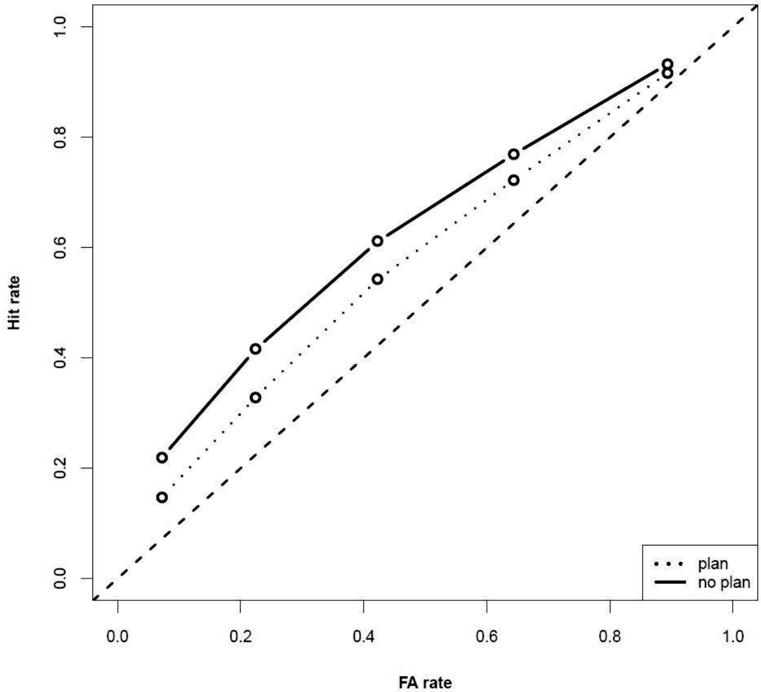

Figure 6. Group ROC curves for Experiment 2. The x-axis represents False Alarm (FA) rate, and the y-axis Hit (H) rate. The 5 circles on each task line correspond to the 5 different scale-splits (from left to right: 1_23456, 12_3456, 123_456, 1234_56, 12345_6. The numbers that appear left of the underscore represent "new"- and the numbers right of the underscore represent "old"- responses). The dashed diagonal line represents chance level. Values in the upper right corner of the graph represent a liberal criterion whereas values in the lower left corner of the graph represent more conservative criteria.

To evaluated decision speed as a function of task (plan, no-plan) and recognition memory value we removed responses equal to or faster than 300 ms from the analyses (6% of the data). A further 4% of the data were eliminated as outliers. Decision speed was entered as the dependent variable in a model that included task (plan and no-plan, fitted as 1 and -1 respectively), recognition memory value (6-point scale centered), list (List 1 and List 2, fitted as -1 and 1 respectively), and the interaction of task with recognition memory value, as fixed factors. The random effects structure included random intercepts for participants and spoken words as well as random slopes (by participants and by spoken words) for task and for recognition memory value and the corresponding interactions.

The model that best fit the decision speed data included significant main effects of recognition memory value ($b = -43.995$, $p < .001$, intercept = 1941.486) with high recognition memory values leading to faster decision than low recognition memory values, whereas task did not contribute significantly to the model fit ($b = 10.615$, $p = .132$, intercept = 1941.486; $\text{Plan}_{\text{mean}}$ =1987 ms, $\text{Plan}_{\text{standard deviation}}$ = 211; $\text{No-Plan}_{\text{mean}}$ =1967 ms, $\text{No-Plan}_{\text{standard deviation}}$ = 223).

An additional analysis evaluated how decision speed for lures related to decision speed for targets in the two tasks. "Lure-old response" was set as the reference level. The model - with task, response and the interaction of task by response as fixed factors and a random intercept by participants and spoken words as well as random slopes for task and response - that best fit the data indicated that for “old” responses, no-plan items were decided on faster than lures ($b_{\text{no-plan old}} = -69.47$, $p < .001$, intercept = 1940.33; $\text{No-Plan-Old}_{\text{mean}}$ =1921 ms, $\text{No-Plan-Old}_{\text{standard deviation}}$ = 201; $\text{Lures-Old}_{\text{mean}}$ =1979 ms, $\text{Lures-Old}_{\text{standard deviation}}$ = 231), whereas plan items did not differ from the lures ($b_{\text{plan old}} = -25.69$, $p = .187$; $\text{Plan-Old}_{\text{mean}}$ =1949 ms, $\text{Plan-Old}_{\text{standard deviation}}$ = 178; $\text{Lures-Old}_{\text{mean}}$ =1979 ms, $\text{Lures-Old}_{\text{standard deviation}}$ = 231). A significant interaction of response with the no-plan task ($b = 84.53$, $p < .001$) revealed that the difference between no-plan items and lures was bigger when labeling and item as “old” (-69 ms) compared to “new” (15 ms). The interaction of response with the plan task was also significant (b = 48.18, p =.033) with the difference between plan items and lures being bigger for “old” (-26) as compared to “new” responses (22 ms). Thus again no-plan hits (“old” no-plan) were decided upon faster than false alarms (“old” lure), whereas a plan hit (“old” plan) was decided on in similar speed to a false alarm (“old” lure).

Pupil Size Data. Figure 7 depicts the average percentage of pupil diameter change (PDC) in the recognition memory test as a function of response in the plan and no-plan task of the exposure phase. An “old” response was accompanied by larger dilation than a “new” response. Two mixed effects models, for mean and peak pupil size respectively were assessed. Both initial models included fixed effects for task (plan and no-plan fitted as -1 and 1 respectively) and recognition memory value (6-point scale centered) as a simple numeric term, but also its quadratic part, in an attempt to model the u-shaped curve of the pupil size data. The interaction term of task by response was also part of the fixed structure. The random structure of the models included random intercepts by participants and spoken words, as well as random slopes for task and response and their interaction by participants and spoken words. The final models included fixed effects of the simple (mean pupil size: b =.0028, p =.001, intercept =.0127; peak pupil size: b =.0046, p <.001, intercept =.0625) and the quadratic term of response (mean pupil size: b =.0030, p <.001; peak pupil size: b =.0050, p <.001) and non-significant fixed effects of task (mean pupil size: b =.0013, p =.165; peak pupil size: b =.0017, p =.155), with random intercepts for participants and spoken words, along with random slopes for task and response by participants and spoken words.

As in Experiment 1, we evaluated whether the encoding pattern for the spoken words presented in the no-plan task of the exposure phase – as measured via pupil dilation - would predict later correct recognition memory value (6-point scale) for these items in the recognition memory test. The two initial models included mean or peak pupil size as a fixed factor and random intercepts by participants and spoken words. Correct recognition of a target audio in the recognition memory test was not linked to larger dilation during encoding in the exposure phase, even though again a trend towards significance arose for mean pupil size (mean pupil size: b = 1.489, p =.057, intercept = 0.466; peak pupil size: b = 0.988, p =.14, intercept = 0.418).

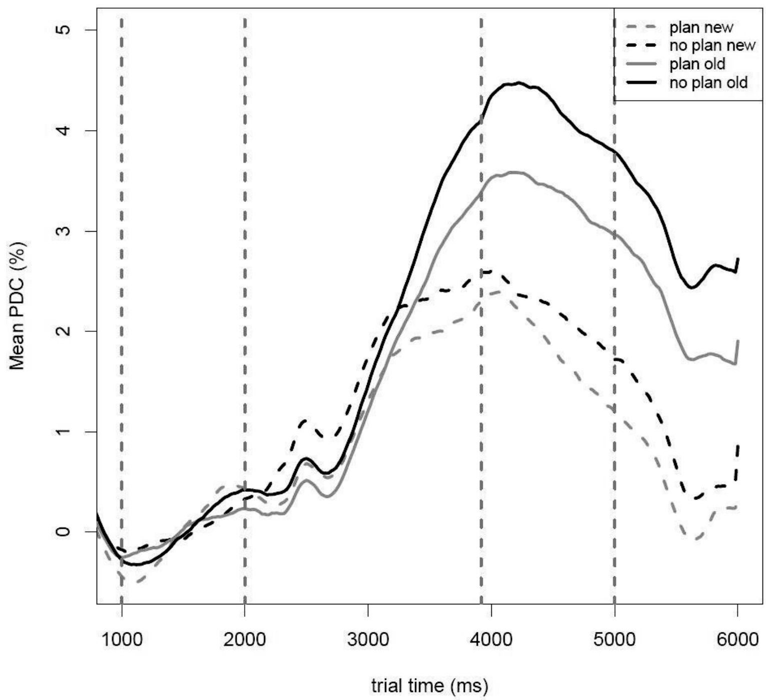

Figure 7. Average percentage of pupil diameter change (PDC) in the recognition memory test, as a function of response (new/old, based on the 3-by-3 scale split) and task (plan, no-plan) in Experiment 2. The x-axis represents trial time in ms. The vertical dashed lines (starting at the leftmost line) indicate baseline (1000 to 2000 ms), scale and audio onset (2000 ms), mean RT and scale offset (5000 ms). The critical time-window during which pupil size was assessed was from 2000 – 5000 ms. Plan and no-plan new are misses, while plan and no-plan old are hits.

Comparison of Experiment 1 and 2

In order to assess how the memory instructions affected the participants' performance, we compare the two experiments on a number of behavioral measures. The same analyses as per experiment were run with experiment as an additional factor. We report only data that concern main effects of experiment or the interaction of experiment with another factor.

Exposure phase. Participants made more naming errors in Experiment 2 (14.54%) than in Experiment 1 (12.24%), but did not differ substantially in the picture naming latencies. Figure 8 depicts error rates for high and low frequency picture-names in the exposure phase of the two experiments. To assess this pattern, a model was fitted to error rates that included experiment (Experiment 1 and Experiment 2, fitted as -1 and 1 respectively), picture LogF (high and low, fitted as -1 and 1 respectively), list (List 1 and List 2, fitted as -1 and 1 respectively) and the interaction between experiment and picture LogF as fixed factors, random intercepts for participants and picture-names, and a random slope for picture LogF by participants. This model revealed significant effects for experiment ($b = .195$, $p = .037$, intercept = -2.775), picture LogF ($b = .784$, $p < .001$), as well as the interaction between experiment and picture LogF ($b = -.160$, $p = .018$). The final model also included list ($b = -.032$, $p = .859$). The interaction between experiment and picture LogF seems to arise from the fact that the difference between high and low frequency pictures was larger in Experiment 1 (13.96%) than in Experiment 2 (10.52%). This difference was due to the fact that participants made more errors in the high frequency pictures of Experiment 2 (9.28%), than in the high frequency pictures of Experiment 1 (5.44%).

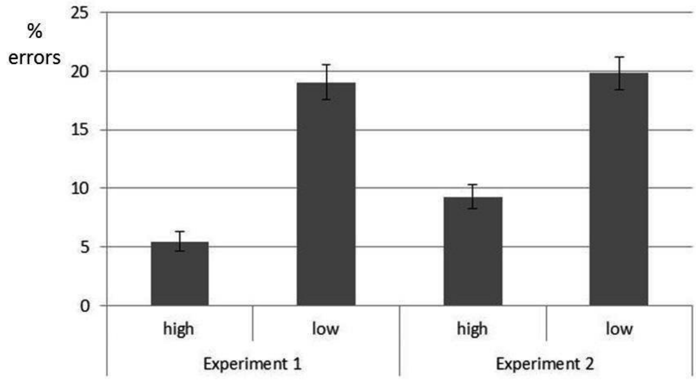

Figure 8. Error rates for high and low frequency picture-names in the exposure phase of the two experiments. Error bars reflect the standard error of the mean.

Comparison of response speed in the two experiments using the same model configuration as for error rates demonstrated that participants displayed similar naming paces across the two experiments (experiment: b = -10.15, p =.596, intercept = 1071.22) and were thus not slowed down by the need to also listen carefully to the spoken words as a result of the expected upcoming recognition memory test in Experiment 2.

Recognition memory test. As a comparison between Figures 3 and 6 shows, the participants of Experiment 2 seem to have outperformed those of Experiment 1 in the recognition memory task. To compare the groups statistically, the sensitivity measures were submitted to two-way mixed Analyses of Variance (ANOVA) by participants. Task (plan, no-plan) was entered as a within-participant factor and experiment as a between participants factor. The analysis revealed main effects of task (F_1 (1, 61) = 43.749, $p <.001$, η^2 =.418; M_{plan} =.207, SE_{plan} =.032, $M_{No\text{-}plan}$ =.433, $SE_{No\text{-}plan}$ =.032) and experiment (F_1 (1, 61) = 9.054, p =.004, η^2 =.129; $M_{Experiment\ 1}$ =.238, $SE_{Experiment\ 1}$ =.039, $M_{Experiment\ 2}$ =.401, $SE_{Experiment\ 2}$ =.037). The interaction between task and experiment was not significant (F_1 (1, 61) =.873, p =.354, η^2 =.014), which demonstrates that the difference in recognition memory accuracy between the plan and the no-plan task was not significantly modulated by the instructions to ignore or pay attention to the spoken words.

Analyses of decision speed did not reveal any main effects or interactions of experiment.

Discussion

In Experiment 2 we provided participants with intentional encoding instructions to test whether the findings of Experiment 1 - inferior memory performance for items encoded during planning speech as compared to items encoded during no planning - could be ascribed to the incidental environment we had created for our participants. Independently of encoding environment, however, auditory targets encoded during planning a single word led to inferior recognition memory accuracy than items encoded during no speech planning. As in Experiment 1 pupil dilation data supported the behavioral patterns. "Old" responses to auditory targets (hits) led to larger dilation than "new" responses (misses) thereby linking pupil dilation during the recognition memory task to memory retrieval processes.

Contrary to our expectations, pupil size in the recognition memory test was not modulated by task. A modulation of pupil size in the recognition memory test by task would have possibly indicated that retrieval was more effortful for heard words that were encoded during planning compared to those encoded during not planning. In considering why we could not pick up such a pattern and in an attempt to explain possible reasons for the discrepancy between our results and those of Otero and colleagues (2011) we focused on some design differences in our experimental setup, which could be of importance. In their study participants heard 30 words per block, while either counting the number of syllables of the word (shallow encoding) or generating a synonym for the word (deep encoding). Every block was immediately followed by the recognition memory test for these 30 targets, which were intermixed with 20 lures. Thus the participants in that study seem to have had an easier task compared to our participants who had to go through all 100 trials of the exposure phase before participating in the 200 trial long recognition memory test.

To assess whether task complexity might have obscured the effect of task, we evaluated pupil size of the first 50 trials out of the 200 trials of the recognition memory test. Testing the same model as for all 200 trials, now on the first 50 trials, revealed significant fixed effects of the simple (mean pupil size: $b = .0033$, $p = .002$, intercept $= .0142$; peak pupil size: $b = .004$, $p = .002$, intercept $= .0588$) and the quadratic term of response (mean pupil size: $b = .0029$, $p < .001$; peak pupil size: $b = .005$, $p < .001$) as well as significant fixed effects of task (mean pupil size: $b = .0040$, $p = .015$; peak pupil size: $b = .0046$, $p = .015$), with random intercepts for participants and spoken words, along with random slopes for task and response by participants and spoken words. Thus intentional encoding revealed a tendency for pupil size in the recognition memory test to be modulated by task, with no-plan auditory targets leading to larger dilation than plan items. Even though this was only the case for the first 50 out of 200 trials in the recognition memory test, it constitutes some evidence that encoding depth did differ between the two tasks, and that this affected retrieval effort in the recognition memory test, at least for the items that were encoded first in the exposure phase and recognition memory test[6]. This suggests that our task was probably more complex than the task employed by Otero and colleagues (2011) and might thereby have obscured any effect of encoding depth in our experiments. Conducting the same analysis for mean and peak pupil size of the last 50 trials of the recognition memory test did not reveal any effect of the simple term of response and of task[7].

Comparing the sensitivity indices in Experiment 1 to those of Experiment 2 revealed that intentional encoding boosted overall memory performance in the recognition memory test, but did not interact with the effect of task. Thus, independently of encoding environment, auditory targets encoded during planning a single word led to inferior recognition memory accuracy than items encoded without concurrent speech planning.

Intentionally dividing attention between listening and planning in Experiment 2 had an impact on participants' naming performance in the exposure phase: Error rates, particularly for high frequency items, were higher compared to Experiment 1. The reasons why the instructions particularly affected performance for high frequency items are unknown; but the pattern of results does indicate that the gain in perceptual processing (here measured as recognition memory performance) in Experiment 2 was accompanied by a performance decrement in the speech production task.

General Discussion

Two experiments investigated the effect that planning a single word had on the perception and memory for a concurrently presented auditory word. In Experiment 1, participants were instructed to ignore the words but they were probed for memory in a surprise recognition test. In Experiment 2 participants were informed about the memory test and were encouraged to divide their attention between word planning and listening. In both experiments, the stimuli were timed in such a way that the participants' word planning - but not the actual articulation - coincided with the presentation of the spoken words. In addition to recording the participants' speech and recognition performance, their pupil size was recorded throughout the experiment.

Memory encoding of heard words is affected by concurrent speech planning.

The experiments confirmed our hypothesis that concurrent speech planning and listening comes with a cost. The participants' ability to detect old words was above chance, but it was significantly better when those words had been heard when participants were not engaged in speech planning than when they were planning a picture name. The cost was evident in recognition memory sensitivity measures as well as in decision speed and pupil dilation patterns.

During the recognition memory test of both experiments the participants' pupils dilated more when they correctly identified an item as old (hit) compared to labeling it as new (miss). This finding is in line with previous work by Papesh and colleagues (2012). Contrary to Papesh and colleagues, however, we only observed a weak tendency for mean pupil size during the exposure phase to predict recognition memory value at test. This difference is probably due to task differences. Participants in the study by Papesh and colleagues always focused on a fixation cross at the exposure phase, with no additional task, while our exposure phase was considerably more complex, both visually (with the presentation of pictures or the meaningless line drawing) and cognitively (dual-tasking).

The current results demonstrate that speech planning affects listeners' ability to form a reliable memory trace of concurrently presented spoken words. Important questions for further research concern the origin of the effects observed here. At present, it is not clear which components of the word recognition processes were most strongly affected by concurrent speech planning. Potential candidates are early auditory processes (see Mattys & Wiget, 2011, for arguments in favor of impoverished encoding of the sensory input under increased cognitive load), mapping of words onto lexical representations, or the generation of episodic memory representations of words. Likewise, it is not clear which aspects of the speech planning processes caused the disruption of the processing or storage of the spoken words. This could be early visual[8]-conceptual and/or subsequent lexical retrieval processes[9]. Future research varying the properties of the spoken words and pictures and the timing of their presentation may elucidate these questions.

Monitoring the cognitive effort of planning while listening via pupil dilation.

During the exposure phase, we observed that planning a single word while listening to a single word led to larger pupil dilation than just listening to a single word. Moreover, planning a low frequency picture name led to larger pupil dilation than planning a high frequency picture name, confirming the relation between TEPR and cognitive demand. TEPR's may prove a useful tool in investigating cognitive effort during linguistic dual-task situations such as turn-taking in interaction, or in general to establish how cognitive load is distributed during listening and speaking in a dialogue situation. Research by Wierda, Rijn, Taatgen, and Martens (2012) showed that pupil dilation can be used to track attention at high temporal resolution.

Focusing attention on planning only (Experiment 1) versus dividing attention between planning and listening (Experiment 2).

The comparison across experiments showed that explicit attention to the spoken words during the exposure phase (Experiment 2) improved recognition performance at test. However, even when participants attended to the spoken words (Experiment 2), planning a picture name still led to a decrement in recognition performance compared to only listening to the words. In the exposure phase, we observed that incidental and intentional encoding also affected naming accuracy differently. That is, participants made more errors in picture naming, especially for high frequency pictures, when dividing their attention between planning and listening (Experiment 2) than when only focusing on planning (Experiment 1). This difference could not be attributed to differences in naming speed in the two experiments.

These two general effects of dividing attention - costs in naming accuracy in the exposure phase and benefits for recognition memory performance in the recognition memory test - reflect the tug of war between concurrent perception and production processes. It appears that there is a limited pool of resources that has to be distributed between the two tasks, and that adding resources to one process (e.g., by paying more attention to spoken words) may lead to a drop in the resources available for the other task (e.g., more picture naming errors). Despite the extended experience people have with concurrent speech planning and listening, the two processes competed with each other.

Implications for conversational settings and future directions.

The present work arose from our interest in understanding the dynamics of dialogue, and especially situations of turn taking. In turn taking, planning to speak and listening have been found to overlap for some amount of time (Bögels et al., 2015; Boiteau et al., 2014; Bosch et al., 2005; Levinson, 2015; Sjerps & Meyer, 2015). These researchers have suggested differences in the precise timing of onset of planning however, where some have argued that planning of one's own response starts as early as possible (Bögels et al., 2018, 2015; Levinson, 2015), while others have argued that planning starts only shortly before the transition of a turn (Boiteau et al., 2014; Sjerps & Meyer, 2015). One potential reason for such discrepancy is that different factors may govern an upcoming speaker’s decision to start planning their speech early or late (see also Barthel, Meyer, & Levinson, 2017). One such factor is the potential influence that planning may have on listening. As of yet, however, it was unclear whether overlap between listening and speech production planning would affect the quality of listening (as evaluated via recognition memory performance). Only a very recent study has addressed this question, using an N400 paradigm, and concluded that an impact of production planning on listening was mainly evident in so called quick responders, whose N400 size effect was smaller than that of slow responders, suggesting that the former shifted their attention away from listening earlier than the latter (Bögels et al., 2018). Given that speaking and listening are highly overlearned activities, one might expect people to be able to carry out these activities in parallel without much interference. Many earlier studies have shown that single word production suffers from the simultaneous presentation of spoken words (e.g. Damian & Martin, 1999; Schriefers et al., 1990). The current study shows that the converse is also true: Listening to words is affected by concurrent speech planning.

A number of aspects about the impact of mutual interference between speech planning and listening on the success of everyday conversations need to be further determined however. In the present study, we investigated the memory for isolated words. In conversation, interlocutors usually care especially about the meaning of sentences (though occasionally, for instance in a legal dispute, memory for the exact wording of utterances is important). On the basis of the present data, we cannot say whether or not grasping the meaning of utterances is affected by concurrent speech planning. According to Bock, Dell, Garnsey, Kramer, & Kubose (2007) comprehension effort usually peaks near the ends of utterances, while the demands of production seem to reach a maximum around or even before the beginning of the utterance. If concurrent speech planning impacts on conversation, then interlocutors might have to adopt a strategy by which they prioritize planning or listening at certain points in the flow of the conversation. This situation is similar to the need to share central resources between conversing and driving as discussed by Becic et al., (2010). Interlocutors can schedule their speech planning processes as they see fit. They might, for instance, opt to begin to plan an utterance while another person is still speaking (but at the cost of listening quality), or they might postpone their planning until the other person has completed their turn. As such, listeners may exert some control over the quality of their comprehension and planning processes that depend on the situation-specific demands. For example, they could aim for a deep understanding of the incoming utterance or for the formulation of an elegant or thoughtful reply (see Ferreira, Bailey & Ferraro, 2002; Ferreira, & Patson, 2007 for discussion on good enough perception, and Swets, Jacovina, Gerrig, &

Gerrig, 2013 for discussion of good enough production). Or they could try to do both, but then at the cost of a long inter-turn interval, which contains the risk of losing one's turn.

More research is needed to elucidate how interlocutors deal with the multiple cognitive demands arising during conversation. Perhaps such an approach may also reveal why, in spite of these demands, conversing is typically experienced as such an easy and pleasurable activity.

Acknowledgments

The authors thank Alastair Smith for his comments on an earlier version of this chapter.

Footnotes

[1] d' rather than some other measure like proportion correct was used as a measure of accuracy, because d' is a robust measure that depends on stimulus parameters and remains roughly invariant when response bias is manipulated. This does not hold for measures like proportion correct.

[2]The plan task could not be evaluated on that since we could not discern what proportion of the dilation was the product of perception and what proportion the product of planning.

[3] The random structure is mentioned once for each dependent measure, unless convergence problems forced us to alter it in any way. If not explicitly stated, the random structure described at the beginning of each section applies to all the models of this dependent measure.

[4] The estimates correspond to ratios as measured from relative pupil size and not to percentage of pupil change.

[5]Note that this is not the classic pupil old/new effect.

[6] Item order was kept constant in the exposure phase and the recognition memory test.

[7] In the same analyses for Experiment 1 we found no effect of task for either the first or the last 50 trials.

[8] It is unlikely that the observed differences can be fully attributed to the fact that there was only a relatively simple picture in the no-plan task (meaningless line drawing) vs. the more variable visual stimuli in the plan task. A reliable difference in pupil dilation was also observed between high and low frequency picture names (which were matched in visual complexity) in the exposure phase, with low frequency picture names leading to larger dilation.

[9] It has been shown that the neuronal infrastructures for speaking and listening show substantial overlap at the lexical level (Menenti, Gierhan, Segaert, and Hagoort, 2011). Lexical retrieval for naming a picture may therefore overlap with lexical retrieval processes for listening to the spoken word and may thus cause interference when engaged concurrently.

[10] The names are adopted from Severen's database.

Chapter 4

Speech planning affects memory for sentence final words

Gerakaki, S., Sjerps, M.J. Speech planning affects memory for sentence final words.

Abstract

When people converse they constantly switch between being a speaker and being a listener. Although subjectively this process unfolds without much effort, it is unclear how well we can listen while we are planning our own speech. In chapter 3 we demonstrated that planning a single word interferes with listening to a single word. This chapter investigates how well listeners can cope with the dual demands of planning single words while listening to whole sentences. In the exposure phase of Experiment 1we manipulated production effort by asking participants to name a picture on half of the trials (plan task) while the other half featured a meaningless line-drawing which was not named (no-plan task). Monitoring pupil dilation during the exposure phase revealed increased pupil size during planning. Listening performance was assessed through a subsequent recognition memory test that evaluated how well participants remembered the sentence-final-words of the exposure phase. The recognition memory results confirmed that sentence-final-words that had been heard while preparing to speak were indeed recognized less often than sentence-final-words heard without concurrent speech planning. We also manipulated the predictability of the sentence final words -by embedding these in either a constraining or a non-constraining sentence frame- to investigate whether predictability may alleviate the observed interference. Surprisingly, however, there was no effect of sentence-final-word predictability on recognition memory. A control experiment (Experiment 2) confirmed that the difference in recognition memory performance between the plan and no-plan task was indeed dependent on differences in planning effort rather than on differences in effort of picture-recognition-processes between the nameable pictures and the meaningless line-drawing. The data presented here

demonstrate that planning to speak interferes with the quality of listening (at least when assessed via recognition memory performance) in sentence context, and links this effect to cognitive effort.

CHAPTER 4: SPEECH PLANNING AFFECTS MEMORY FOR SENTENCE FINAL WORDS

Introduction

In conversation interlocutors constantly switch between being a speaker and being a listener. Because pauses between interlocutor's turns are typically short, and turns sometimes even overlap in time (e.g. Beattie & Barnard, 1979; Bosch, Oostdijk, & Boves, 2005; Stivers et al., 2009), researchers have proposed that people regularly begin to prepare their verbal response while still listening to the other person (Bögels et al., 2018, 2015; Boiteau et al., 2014; Sjerps & Meyer, 2015). Indeed, chapter 2 of the current thesis demonstrated that participants may start to prepare their response while still listening, when presented with a constant sentence in which only the words "left" and "right" and their position in the sentence (early-late) are alternated. It is unclear, however, whether such overlap between speech planning and listening may have a negative effect on either of these processes.

In chapter 3, we demonstrated that planning single words while listening to another word does have a negative impact on how well people remember what they heard. In more typical conversation, however, people listen to more than single words, and may be able to overcome potential interference by relying on constraining information in the earlier parts of their interlocutors' sentences. For example, it has been demonstrated that predictable words are processed faster than less predictable ones (see for example Altmann & Kamide, 1999; Kliegl, Nuthmann, & Engbert, 2006; Traxler & Foss, 2000). Predictable words may thus have a processing advantage over less predictable ones, which can arise because predictable words are indeed predicted or because they can be more readily integrated into the preceding sentence context (e.g. Hintz, Meyer, & Huettig, 2016 & 2017; Huettig, 2015). It is unclear, however, whether such facilitative effects may also help to overcome the potential interference that arises because of production-related processes.

Here we investigated 1) whether listening to sentences is affected by concurrent speech planning and 2) whether any observed effects of planning speech on listening to speech can be mediated by the predictability of the sentence-final-word. To address the first question we manipulated production effort by presenting participants with either nameable pictures (plan task) or a meaningless line-drawing (no-plan task). In addition, among the nameable pictures we presented objects with frequent picture-names (low-effort retrieval) or objects with infrequent picture-names (high-effort retrieval). To address the second question we manipulated the predictability of heard words by presenting listeners with sentence frames that were constraining- or non-constraining- with regard to the sentence-final-word. The use of spoken sentences in the listening task not only aimed at creating the context for predicting the sentence-final-word but also at moving our design a step closer to more ecological conversational settings compared to the experiments reported in chapter 3. That is, we investigated whether the recognition memory decrement that we had observed in single word planning while listening (chapter 3) would persist when participants listen to full sentences and name pictures presented concurrently with the sentence-final word.

Participants thus listened to sentences where the sentence-final-word was either predictable or not predictable given the preceding context: (1) De boer melkte de *koe* (The farmer milked the *cow*) or (2) Het kind tekende een *koe* (The child drew a *cow*). In addition, a line drawing of an unrelated picture or the meaningless line drawing appeared time-locked to the onset of the last word of the sentence. The participants were asked to listen to the sentences attentively and, as in the study described in chapter 3, name the pictures as quickly as possible (for nameable pictures), or refrain from responding (for the meaningless line drawing). As before, the trials featuring nameable line drawings constituted the plan task, whereas the trials featuring the meaningless line drawing constituted the no-plan task.

We had a number of main predictions. First, on the listening side, we expected that target words would be processed less efficiently -and hence remembered less well- when they had been heard when participants were engaged in the planning of naming a picture. Second, we predicted that target words would be processed more efficiently -and hence remembered better- when they were predictable given their preceding context than when they could not be predicted from prior context (see also Besson, Kutas, & Van Petten, 1992; Neville, Kutas, Chesney, & Schmidt, 1986; Olichney et al., 2000, showing better memory for semantically congruous than incongruous words in recall- and recognition-memory tests). Moreover, a memory advantage for predicted, compared to not predicted, words might be most pronounced in the plan-task, where processing resources are reduced due to the concurrent speech planning task. Third, on the production side, we expected that picture naming latencies may be shorter in the context of trials with predictable sentence-final-words than with unpredictable sentence-final-words (i.e., an influence of ease-of-listening on production performance) as production-perception interference may operate in both directions.

To relate ease of recognition during the recognition memory test to processing load during exposure, we measured reaction times and error proportions of picture naming in the exposure phase, along with continuous pupil dilation in both phases. Task-evoked pupillary responses (TEPR) are changes in pupil size that have been shown to reflect processing demands (e.g. Beatty & Lucero-Wagoner, 2000; Kahneman, 1973; Laeng, Sirois, & Gredeback, 2012). The effect of processing demands on pupil size has, for example, been demonstrated by increased pupil sizes during the simultaneous interpreting of difficult as compared to easy words (Hyönä, Tommola and Alaja, 1995). These combined measures should provide us with a reliable measure of overall processing demands during encoding in the exposure phase and ease of recognition during the recognition memory test.

Finally, as noted, the plan task of Experiment 1 featured a set of different pictures, whereas the no-plan task featured a single meaningless line drawing. Experiment 2 was aimed at testing whether any memory advantage found in Experiment 1 for sentence-final-words heard in the no-plan task was due to the difference in the visual stimuli rather than the difference in cognitive task demands between the plan and no-plan conditions.

Experiment 1

Method

Overview. Participants went through an exposure phase and a recognition memory test. In the exposure phase, they listened to sentences which either did or did not render the last word highly predictable. At the onset of the sentence-final-word either a picture appeared on screen, which participants had to name (plan task) or a meaningless line drawing to which participants did not have to respond in any way (no-plan task). Participants were instructed to name the pictures as quickly as possible, but to also listen carefully to the sentences because they would later "have to do something with these sentences".

In the test phase participants completed a recognition memory task. They were presented with the last words of the sentences which they had heard during the exposure phase (hereafter sentence-final-words), intermixed with an equal number of lures. Participants were asked to indicate whether or not they had heard each word in the exposure phase.

Following the design of chapter 3, the "number-letter" task was used as a filler task between the exposure and the recognition memory test (Miyake et al., 2000; Rogers & Monsell, 1995). The results of the filler task were not evaluated in any way and are thus not further reported here.

Participants. 43 participants were recruited from the participant pool of the Max Planck Institute for Psycholinguistics, Nijmegen. They received a monetary reward for their participation. Eleven participants had to be excluded due to either not conforming to task instructions (n = 2), technical errors (n = 2), or applying a too conservative or too liberal lambda criterion upon deciding in the recognition memory test (n = 7; lambda criterion values were kept between 0 and 1.5). A high value indicates a conservative criterion resulting in few false alarms (saying "new" all the time), while a low value indicates a liberal criterion resulting in many false alarms (saying "old" all the time; for more on the lambda criterion see Wickens, 2001, pages 12 to 14). All participants were adult native speakers of Dutch (mean age = 22.3 years, 28 female), had normal or corrected-to-normal vision, no speech or hearing problems, and were not diagnosed with dyslexia. None had participated in the experiments of Chapter 3. The Ethics Board of the Social Sciences Faculty of the Radboud University Nijmegen gave its consent for conducting the study.

Materials and Design. To generate the exposure materials 50 sentence pairs were selected from the materials used in Piai, Roelofs, and Maris (2014). Each pair featured the same sentence-final word, but in one member of the pair the final word was highly predictable from the preceding context, and in the other member it was not predictable (see Appendix E). For example, the sentence-final-word "anker" ("*anchor*") was once embedded in the constraining sentence frame "Het schip werd stevig vastgelegd met een anker" ("*The ship is firmly established with an anchor*") and once in the non-constraining sentence frame " Dat zware metalen stuk daar is een anker" ("*This heavy metal thing there is an anchor*"). Sentence-final-words of the predictable condition (constraining sentences) had a high cloze probability, and sentence-final-words of the unpredictable condition (non-constraining sentences) had a low cloze probability (for further details see Piai et al. 2014). In order to avoid that participants inferred anything about the purpose of the study, we used 4 additional sentences, rather than two sentence pairs as practice material.

In addition to the cloze probabilities, we determined the plausibility of the sentences. To this end, we asked 20 participants to read the sentences and rate their plausibility on a 7-

point scale (1=not plausible at all; 7 = very plausible). As expected, predictable sentences received higher plausibility rating than unpredictable ones ($M_{predictable} = 6.32$, $SE_{predictable} = .54$; $M_{unpredictable} = 5.08$, $SE_{unpredictable} = 1$; $t(198) = -10.85$, $p < .0001$). Note that for the present purposes, the confound of plausibility and predictability is not problematic, since we were not specifically interested in predictability, but merely aimed to generate materials where the sentence-final word was more or less easy to process.

The sentences were recorded in a sound shielded booth by a female native speaker of Dutch. In the recording list the sentences were intermixed, with the predictable and unpredictable version of a given sentence-final-word never being less than five sentences apart. Rather than splicing the same recording of the sentence-final-word onto the predictable and unpredictable sentence frame, we used the recordings of the whole sentences. The sentences of the predictable condition were on average longer in duration (2492 ms) than those of the unpredictable condition (2344 ms; $t(398) = -3.124$, $p = .002$), but the duration of the sentence-final-word did not differ between conditions ($M_{predictable} = 493$ ms, $SE_{predictable} = 113$; $M_{unpredictable} = 503$ ms, $SE_{unpredictable} = 109$; $t(398) = .919$, $p = .358$).

For plan trials, each sentence pair was combined with a picture, whose name was semantically and phonologically unrelated to the sentence-final word. We selected 52 line-drawings from the data base provided by Severens, Van Lommel, Ratinckx, and Hartsuiker (2005). Two of the line-drawings were used on practice trials. As in the study reported in chapter 3, for half of the pictures name frequency was high and for the other half low (high- vs low-frequency: $t(35.989) = 12.0548$, $p < .001$). High and low frequency picture names were matched in name agreement (H-statistic), age of acquisition, length in number of syllables and number of phonemes, visual complexity (as specified by Bates et al, 2003) and luminosity -a measure of visual brightness- (see Table 1 and Appendix G).

The nameable pictures were presented as black line drawings on a white background, scaled to fit into frames of 300 by 300 pixels (4.5° degrees of visual angle from the participant's position). For no-plan trials, each sentence pair was presented together with a meaningless line drawing (see figure 1 of chapter 3).

Table 1

Characteristics (means with standard deviations in parentheses) of high and low frequency picture names.

	Frequency	
	high	Low
log frequency	0.80 (0.26)	0.062 (0.12)
name agreement (H-statistic)	0.71 (0.69)	0.77 (0.45)
Age of Acquisition (AoA)	5.63 (0.90)	7.30 (1.10)
number of syllables	2.52 (0.77)	2.75 (0.74)
number of phonemes	6.92 (1.98)	7.41 (1.69)
visual complexity	17327 (11896)	17296 (6769)
luminosity	237.46 (14.40)	237.04 (8.80)

Note. The luminosity value listed here refers to the to be named pictures. The line drawing used for the no-plan task had a luminosity value of 238.68.

In sum, four versions were created of each trial-item (combination of visual and auditory stimulus): The final word was, or was not, highly predictable, and it was paired with a visual stimulus that was either a picture to be named or the meaningless line drawing. These four versions of the trial-items constituted the materials for the corresponding four experimental conditions and were randomly assigned to the four experimental lists of the exposure phase. As a result all lists included the same sentence-final-words, but the conditions under which these were heard differed from list to list. The order of the items (meaning the sentence-final-words) was kept constant in all four lists. Participants were randomly assigned to the experimental lists.

The materials for the recognition memory test were the 100 sentence-final words of the exposure phase (see appendix F for a full listing of the sentence-final-words) and an equal number of lures (see appendix H for a full listing of the lures), which were drawn from the CELEX database (Baayen, Piepenbrock & van Rijn, 1993). Targets (old items) and lures (new items) were matched for spoken duration; lemma frequency and number of phonemes (see Table 2 and Appendices C & D). The words were produced by the same speaker who had produced the sentence materials. The speaker did not know whether a given word was a target or a lure. During the actual experiment participants had to decide for each spoken word whether they had heard that word in the exposure phase or not. For this task, participants were given a button box and had to press as quickly as possible the "new" or "old" button. That is, we used only 2 response classes instead of the six-point scale used in the recognition memory test in chapter 3. The results of chapter 3 demonstrated that a middle split (i.e., responses 1, 2, 3 as “new” and 4, 5, 6 as “old”) was sensitive enough to observe differences in memory performance. "New" and "old" response buttons were counterbalanced (left-right, right-left) across participants. The order of the targets was kept constant between the exposure phase and the recognition memory test. Zero to four lures intervened between successive targets. All participants heard the same test items. Two more lures were added to be used in the practice of the recognition memory test, which also included two of the practice items from the exposure phase.

Table 2

Average spoken durations, log Frequency and number of phonemes, of targets and lures with standard deviations in parentheses.

	Targets	Lures
Spoken duration (ms)	587 (133)	621 (136)
log frequency	2.44 (1.24)	2.28 (1.20)
number of phonemes	4.61 (1.63)	4.89 (1.76)

Apparatus. The same apparatus was used as in the study described in chapter 3. The only difference was that in the recognition memory test participants used a button box with two buttons to respond.

Procedure. Chin and forehead rests of the eye-tracker were adjusted for each participant at the beginning of the session. The written instructions for the exposure phase appeared on screen. Participants had to name all pictures correctly and as quickly as possible, while also listening to the sentences. When the participant indicated that they had understood the instructions, calibration of the eye tracker started using a nine-point calibration procedure.

Every trial in the exposure phase started with the presentation of a fixation cross for 3000 ms followed by the onset of the spoken sentence. Time-locked to the onset of the sentence-final word either a picture or the meaningless line-drawing was presented and remained on screen for 250 ms. Such brief exposure to the visual stimulus was meant to encourage fast speech planning. The visual stimulus was replaced by a fixation cross for 1750 ms. A blank interval of 500 ms followed. The long exposure to the fixation cross at trial onset (3000 ms) and a blank interval at trial offset (500 ms) were necessary to allow for pupil size to return to baseline between trials. Reaction Times (RTs) were measured from picture onset on (note that picture onset was synchronous to the onset of the sentence-final-word).

The experiment started with a four-trial practice session, two of which featured pictures that had to be named. Upon completing the practice trials participants could ask any clarification questions regarding the task. When all questions had been answered the main part of the experiment began. A total of 100 trials, separated in four blocks of 25 trials each made up the exposure phase session. There were short breaks to rest between the blocks and at the beginning of each block a drift correction was performed. If necessary the eye tracker was recalibrated. The exposure phase took approximately 18 minutes to complete. Then the participants completed the number-letter task mentioned in chapter 3. This took approximately 20 minutes.

In the recognition memory test, the same procedure was used as described in chapter 3. The only difference was that participants were now asked to make a binary decision (new/old), instead of choosing a number out of the six-point scale (see above). At the beginning of each trial participants saw a fixation cross for 2000 ms; then the single word was heard. Participants had three seconds to make a decision during which a fixation cross was displayed on screen. A blank interval of one second concluded the trial. The auditory word offset served as the timepoint from which on decision latencies were measured.Two hundred trials, split up into eight test blocks of 25 items each, made up the material of the recognition memory test. This phase started with four practice items, two of which were sentence-final-words heard during the practice session in the exposure phase. At the end of each block the participant could rest. In total, the recognition memory test took approximately 25 minutes.

Recording and analysis of pupil size. Pupil size was recorded throughout the experiment and analysed as described in chapter 3. In addition to all previous analysis steps the signal was also smoothed by using a simple moving average algorithm (the moving window was six samples long). For the exposure phase the critical time window for which mean and peak pupil size was evaluated began at the average sentence-final-word offset and lasted for 2000 ms. For the recognition memory test the critical time window began at the word offset and lasted for 3000 ms.

Results

Exposure phase

Behavioral Data. Responses to the pictures on plan-trials were categorized as correct or incorrect. Incorrect responses included other -but still correct- naming of the picture (e.g. "globe" instead of "wereldbol", which both refer to the Earth Globe), incorrect naming (e.g. "vergiet" (= colander) instead of "trechter" (= funnel), categorical response (e.g. saying "vogel" (=bird) instead of "pelikaan" (= pelican), hesitations, mispronunciations and no response. We also monitored whether participants indeed said nothing during no-plan trials.

The error rates for the naming task were evaluated with generalized linear mixed-effects regression models in R (version 2.14.2; The R foundation for statistical computing;

lme4 package; Bates et al., 2015). The models were fitted with binomial distributions (Jaeger, 2008). A design-driven approach was adopted and the random structure of the models was determined following the suggestions of Barr, Levy, Scheepers, and Tily (2013). To evaluate interactions we used a backward elimination procedure that examined whether including a specific interaction in the model improved model fit. Step by step, interactions that turned out to be non-significant were left out from the equation. All predictors justified by the design were kept in the models. In addition, a null model that included only the random structure of the models was established. Model comparison was made using log-likelihood ratio tests. Statistics of the optimal models are reported.

Error rates were evaluated with a model including the factors picture LogF (high and low; fitted as 1 and -1, respectively), predictability (predictable and unpredictable, fitted as 1 and -1 respectively), plausibility (centered) and experimental list (centered) as fixed effects, and intercepts by participants and picture-names, and slopes for picture LogF and predictability by participants[1] as the random effects. Only picture LogF was a significant factor in predicting error rates (b = -.45, p =.02, intercept = -2.7). Participants made more errors when naming low frequency (25.4 %) compared to high frequency picture names (17.6 %).

Before carrying out the analyses of picture naming latencies, we also removed responses faster than 300 ms from the analyses (0.4 % of the data), as well as data points with a value further than two standard deviations away from the participant mean (5.5%). We also removed responses slower than 2000 ms (0.7 % of the data). This was done because participants had a maximum of 2000 ms to respond.

Response speed was evaluated in a model that included picture LogF (high and low; fitted as 1 and -1, respectively), predictability (predictable and unpredictable, fitted as 1 and -1 respectively), plausibility (centered) and experimental list (centered) as fixed factors. The random structure of the model was identical to that of the error rates model. The model that best fit the response speed data included a significant main effect of picture LogF ($b = -37.6$, $p = .01$, intercept = 1042) but not of predictability. The interaction of picture LogF and predictability significantly contributed to the model fit ($b = 12.07$, $p = .013$). The factors plausibility and list were also part of the final model, but did not significantly contribute to modeling the response speed data. The main effect of picture LogF was due to high frequency pictures being named faster (994 ms) than low frequency ones (1064 ms). Separate analyses were carried out for the high and low LogF picture names to evaluate the interaction. Predictability, plausibility and experimental list were included as fixed factors, with random intercepts for participants and picture-names and a random slope for predictability by participants only. Naming a low frequency picture was faster when having just heard a predictable sentence-final-word (1047 ms) compared to when having just heard an unpredictable sentence-final-word (1082 ms, $b = -19.12$, $p = .004$, intercept = 1079.68). There was no such effect for high frequency pictures ($b = .083$, $p = .993$, intercept = 1005.46, $M_{predictable} = 999$ ms, $M_{unpredictable} = 989$ ms). Although subtle, these findings demonstrate that production performance can be negatively affected when listening effort increases.

Pupil Size Data. Mean and peak pupil sizes in the plan and no-plan tasks were evaluated to test whether planning a picture-name while listening to a sentence is more cognitively demanding than only listening to a sentence. The initial model included fixed effects for task (plan and no-plan, fitted as -1 and 1 respectively), plausibility (centered) and list (centered) but also for predictability (predictable and unpredictable, fitted as 1 and -1 respectively) and the interaction of task with predictability. The random effects structure contained random intercepts by participants and picture-names, with a random slope of task and predictability by participants. The optimal model included a significant effect of task only ($b_{mean} = -0.038$, p =.001, intercept =.041; b_{peak} = -0.45, p < 0.001, intercept =.096). As depicted in figure 1, planning speech while listening to speech induced larger pupil dilation (Plan mean =$.01943^2$, Plan peak =.1186) than listening (No-Plan mean =.0069, No-Plan peak =.0525).

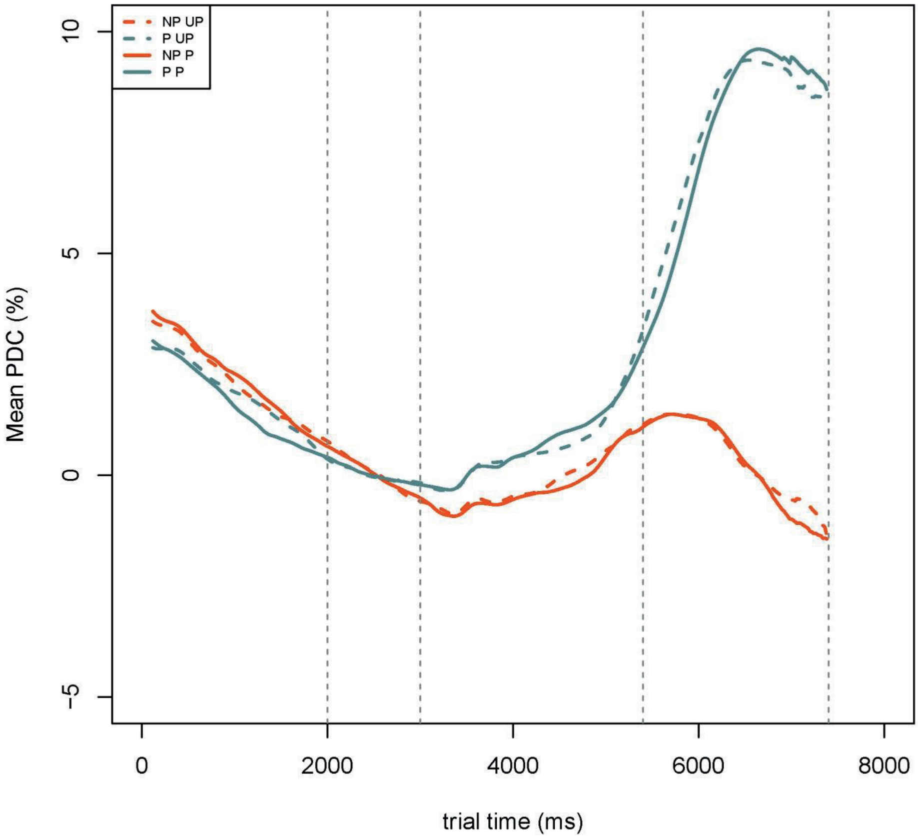

Figure 1. Average percentage of pupil diameter change (PDC) as a function of task (NP = no-plan, P = plan) and predictability (UP = unpredictable, P = predictable) in the exposure phase for experiment 1. The vertical doted lines stand for pre-baseline (up to 2000 ms), baseline (2000 to 3000ms), average offset of audio (5400 ms), and average onset of blank (7400 ms).

The cognitive demand invoked by planning while listening was further evaluated by assessing whether picture LogF modulated the amount of dilation (i.e., analyzing only the plan trials). Two models were tested; one model for mean and one for peak pupil dilation. In both initial models picture LogF (high and low, fitted as 1 and -1 respectively), plausibility (centered) and list (centered) were the fixed factors, with random intercepts for participants and picture-names. Picture LogF was a significant factor in both the mean and peak pupil size model (b_{mean} = -0.008, p =.012, intercept =.079; b_{peak} = -0.011, p =.002, intercept =.141) with low frequency picture names leading to larger dilation (low LogF mean =.088, low LogF peak =.152) than high frequency picture names (high LogF mean =.070, high LogF peak =.129). This shows that the pupil-dilation measures were sensitive to cognitively relevant linguistic aspects of the pictures and that planning the picture-names is associated with increases in general cognitive demands.

Recognition memory test

Behavioral Data. To evaluate whether this increase in general cognitive demands interfered with listening, we assessed recognition memory accuracy as a function of encoding task (plan, no-plan) and predictability. The mean percentages correct were 64.8% and 65.2 % for plan predictable and plan unpredictable, and 66.9 % and 67.5 % for no-plan predictable and no-plan unpredictable, respectively. For the analyses, recognition memory accuracy was quantified by computing the d' (sensitivity signal detection index[3], see Macmillan & Creelman, 2005) for each of the 32 participants in each of the four conditions (i.e., 32 by 4 cells). To this end, responses to target items (i.e. items that were presented during the exposure phase) were categorized as hits (correctly identified as "old", meaning that they were presented during the exposure phase) or misses (incorrectly identified as "new", meaning that they were not presented during the exposure phase) and responses to lure items (i.e. items that had not been presented during the exposure phase) were categorized as correct rejections (correctly identified as "new", meaning that they were not presented during the exposure phase) or false alarms (incorrectly identified as "old", meaning that they were presented during the exposure phase). Figure 2 depicts mean d' values by task and predictability. Items of the no-plan task were remembered better ($M_{No\ plan}$ =.865, $SE_{No\ plan}$ =.014) than items of the plan task (M_{plan} =.55, SE_{plan} =.013). A two-way (plan, no-plan; predictable-unpredictable) repeated measures ANOVA by participants showed a significant effect of task (F_1 (1, 31) = 38.313, $p <.001$, η^2 =.553) but not of predictability (F_1 (1, 31) = 2.278, $p =.141$, η^2 =.068). The interaction of task with predictability was not significant (F_1 (1, 31) =.090, $p =.767$, η^2 =.003).

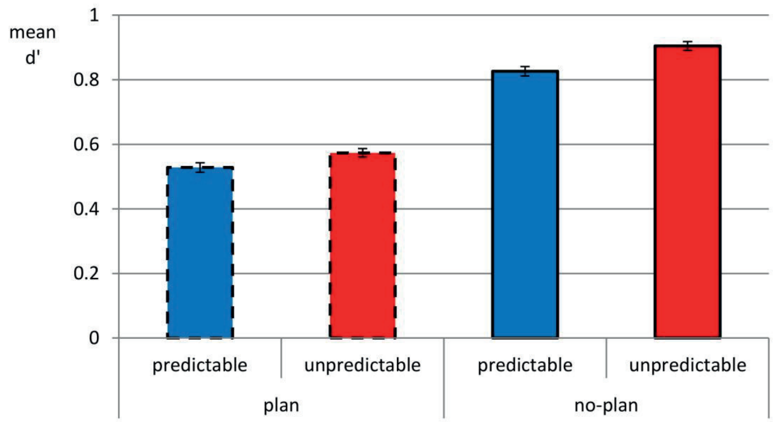

Figure 2. Mean d' value by task (plan, no-plan) and predictability in the recognition memory test of Experiment 1. Error bars reflect the standard error of the mean.

For the analysis of decision speed (Figure 3) we only analysed items that had been presented in the exposure phase (i.e., only targets and not the lures). Outliers were set to two standard deviations from the participant mean (4.23 % of the data excluded). The initial model included task (plan and no-plan, fitted as 1 and -1 respectively), predictability (predictable and unpredictable fitted as 1 and -1 respectively), response (new and old, fitted as -1 and 1 respectively; since only the targets were analysed here, this distinction equals "failed to recognize target", thus a miss in signal detection theory terms, and "succeeded in recognizing target", thus a hit in signal detection theory terms, respectively), plausibility (centered), experimental list (centered) and the interaction of task with predictability and response as fixed factors, a random intercept by participants and spoken word, and a random slope of task and predictability by participants and by spoken words (and the corresponding interactions).

The best fitting model included significant main effects of response (b = -55.36, p <.0001, intercept = 973.94) and a significant three-way interaction of task, predictability and response (b = -20.8, p =.007). Given that such an interaction implies that the two-way interaction differs among the various levels of the 3rd factor, we ran two additional models; one for the level "old" and for the level "new" of the factor "response". Task, predictability, plausibility, and list as well as the interaction of task with predictability were entered as fixed factors in the initial model. The random structure was identical to the model used for evaluating the decision speed overall. The analysis showed a significant contribution of task (plan, no-plan) only when labeling an item as "old" (meaning, when the item was a hit; b = 22.53, p =.035, intercept = 935.83). In particular, no-plan items were labeled "old" significantly faster (895 ms) than plan items (938 ms). No effects were observed for items that were labelled as “new” (meaning when the item was a miss; b = - 3.91, p =.748, intercept = 1050.12). No further significant effects were observed.

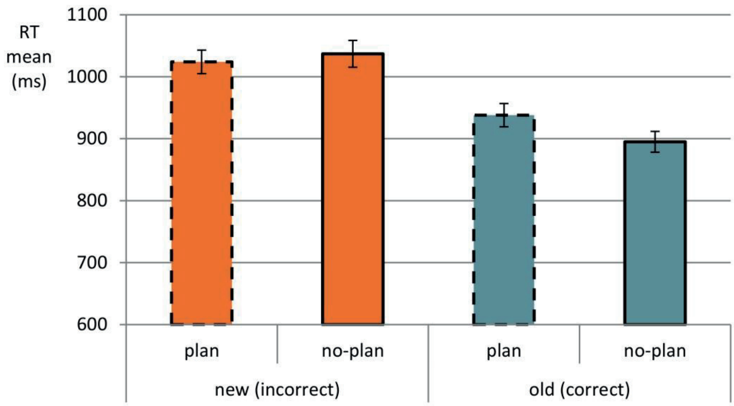

Figure 3. Mean decision speed (ms) for the recognition memory test of Experiment 1 as a function of response and task. The auditory word offset served as the timepoint from which on decision latencies were measured.Error bars reflect the standard error of the mean.

Pupil Size Data. Finally, we evaluated how mean and peak pupil size was modulated by response (new and old, fitted as -1 and 1 respectively, task (plan and no-plan, fitted as -1 and 1 respectively), predictability (predictable and unpredictable, fitted as 1 and -1 respectively), plausibility (centered), experimental list (centered) and the interaction of response, task and predictability. The random structure of the tested models included random intercepts by subjects and by spoken words and a random slope for task and predictability by subjects and by spoken words. Again only target items and not lures were considered for this analysis. Only response contributed significantly to the final models (mean pupil size model: $b = .007$, $p < .0001$, intercept =.044, peak pupil size model: $b = .007$, $p < .0001$, intercept =.106). Thus pupil size was only modulated by whether the participant correctly identified an item as "old" or "new", with "old" responses leading to larger dilation than "new" responses.

Discussion

In this experiment we demonstrated that listening to sentences while preparing to name a picture can negatively affect memory performance for the sentence-final words. That is, when sentence final words had been heard while participants were planning to name a picture, these words were generally recognized less often in the recognition memory test than when the words were heard without concurrent speech planning. In addition, we could relate the recognition memory performance to processing difficulties in the exposure phase through two indices of processing load, reaction time and pupil dilation.

Before discussing these effects in more detail, however, it is important to consider a potential caveat of the influence of planning on subsequent memory. The plan task of Experiment 1 featured a set of different pictures, whereas the no-plan task featured a single meaningless line drawing. Regardless of whether listeners engaged in planning or not, this difference in the novelty/variability of the visual stimuli across the plan and no-plan tasks may have caused differences in the processing demands between these conditions.

In a second experiment we address this concern by replicating Experiment 1 with the crucial difference that listeners were instructed to look at the screen and listen to the sentences passively. That is, in Experiment 2 we only manipulate the visual properties of the stimuli without manipulating speech planning per se, thereby controlling for the effect of stimulus variability per se. If the memory decrement observed in Experiment 1 was caused by the cognitive demands associated with the visual stimuli (instead of cognitive demands of the planning), we should observe the same effects as in Experiment 1. On the other hand, if the memory decrement in Experiment 1 was a result of the planning itself, no effects should be observed in Experiment 2.

Experiment 2

Method

Participants. 38 participants from the Max Planck Institute for Psycholinguistics participant pool took part in the study. Five participants had to be excluded from the analysis, due to maintaining a too conservative or liberal lambda criterion in decision making during the recognition memory test (participants with lambda values below 0 or above 1.5 were excluded; Wickens, 2001), and one for not conforming to task requirements. All were native speakers of Dutch and had normal or corrected-to-normal vision (mean age = 22.7years, 24 female). None of them reported a speech or hearing problem, and none had been diagnosed with dyslexia. Participants gave written consent on their participation and received monetary compensation. As before, participants for which more than 25% of the pupil data trials (including data loss due to erroneous or late responses) in an experimental condition had to be removed were excluded from the analyses. As a result, two more participants were excluded from analyses in the exposure phase and two from analyses in the recognition memory test.

Materials. Materials were identical to Experiment 1.

Apparatus. The apparatus was identical to Experiment 1.

Procedure. The procedure was identical to Experiment 1 with the exception of the instructions. Participants were now instructed to listen to the sentences and only passively view the pictures (that is, they had to keep their eyes on the screen). Thus pictures were not named in this study.

Recording and analysis of pupil size. The recording and analysis protocol were identical to Experiment 1.

Results

We tested the same models as in Experiment 1. Note, however, that to remain consistent across the two experiments we will refer to a plan-control and no-plan-control task for this experiment, though participants never planned any speech for Experiment 2. The pictures, however, were the same as in Experiment 1 such that they were either nameable pictures (in the plan-control task) or the meaningless line drawing (in the no-plan-control task).

Exposure phase

Since this was a passive listening task, there were no behavioural data to analyse.

Pupil Size Data. Mean and peak pupil size did not differ significantly between plan-control and no-plan-control items (mean pupil size model: $b = .003$, $p = .213$, intercept $= .011$; peak pupil size model: $b = .002$, $p = .527$, intercept $= .056$). Furthermore, predictability of the sentence-final-word also did not affect pupil size (mean pupil size model: $b = .0003$, $p = .881$, peak pupil size model: $b = .001$, $p = .608$). Figure 4 depicts the mean pupil dilation change (mean PDC) as a function of task and predictability.

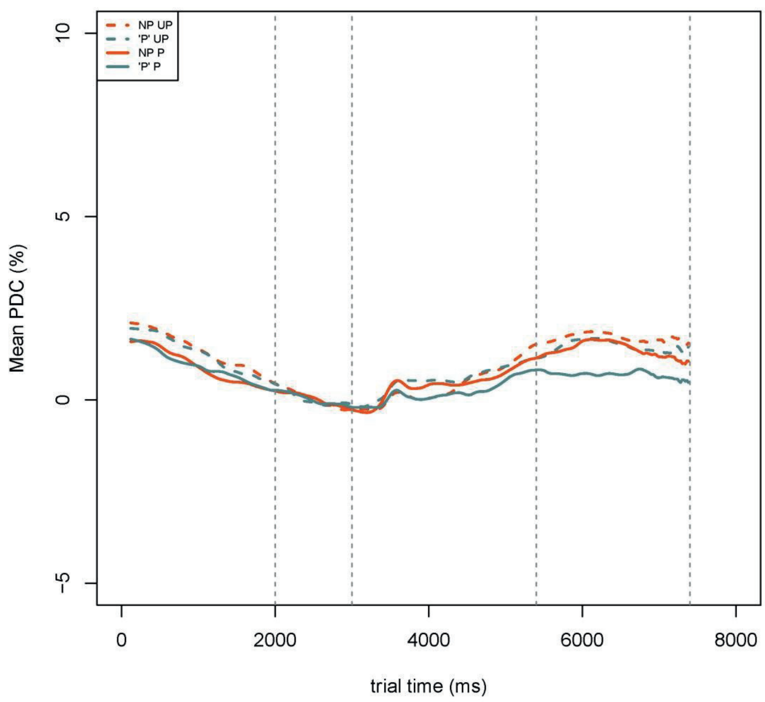

Figure 4. Average percentage of pupil diameter change (PDC) as a function of task (NP = no-plan-control, 'P' = plan-control; note that there was no actual task involved in Experiment 2) and predictability (UP = unpredictable, P = predictable) in the exposure phase for experiment 2. The vertical doted lines stand for pre-baseline (up to 2000 ms), baseline (2000 to 3000ms), average offset of audio (5400 ms), and average onset of blank (7400 ms).

Figure 5 depicts mean pupil dilation change as a function of picture Log F. No significant impact of the factor picture Log Frequency was found on mean or peak pupil size (mean pupil size model: b =.0008, p =.695, intercept =.01; peak pupil size model: b =.0002, p =.910, intercept =.055). Thus, when no naming is involved, low frequency picture-names do not induce larger pupil dilation than high frequency picture-names, thereby demonstrating that the difference in pupil size in Experiment 1 was indeed due to differing planning effort between high and low frequency picture-names and not due to lower-level visual features.

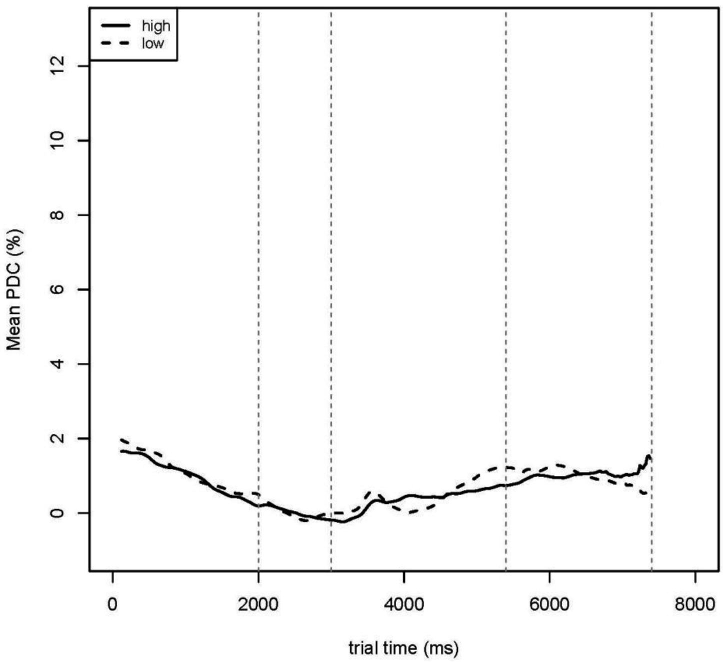

Figure 5. Average percentage of pupil diameter change (PDC) as a function of picture LogF in the plan-control task of the exposure phase for Experiment 2. Note that there was no actual task involved in Experiment 2.The x-axis represents trial time in ms. The vertical doted lines indicate pre-baseline (up to 2000 ms), baseline (2000 to 3000ms), average offset of audio (5400 ms), and average onset of blank (7400 ms).

Recognition memory test

Behavioral Data. Figure 6 depicts mean d' values by task and predictability. A two-way (plan-control, no-plan-control; predictable, unpredictable) repeated measures ANOVA revealed no significant effect of Task (F_1 (1, 31) =.026, *p =.874*, η^2 =.001) nor of Predictability (F_1 (1, 31) = 1.976, *p =.170*, η^2 =.060). Surprisingly, however, the interaction of task with predictability was significant (F_1 (1, 31) = 5.107, *p =.031*, η^2 =.141). Figure 6 indicates that this interaction seems to originate from the fact that sentence-final-words that had been unpredictable were remembered better than those that were predictable, but only so for plan-control trials (i.e., trials with a nameable picture, as compared to the no-plan-control trials that featured the meaningless line drawing). Simple effects dependent t-tests for plan-control and no-plan-control trials confirmed that indeed among plan-control items participants recognized the unpredictable words more often (M = 1.0631, SE =.074) than predictable ones (M =.89, SE =.075, t(31) = -2.308, p =.028). This was not the case among no-plan-control items ($M_{unpredictable}$ =.961, SE =.080; $M_{predictable}$ =.978, SE =.075; t(31) =.280, p =.781). The mean percentage correct was 70.5% and 71.7 % for plan-control predictable and plan-control unpredictable, and 71.1 % and 70.9 % for no-plan-control predictable and no-plan-control unpredictable, respectively.

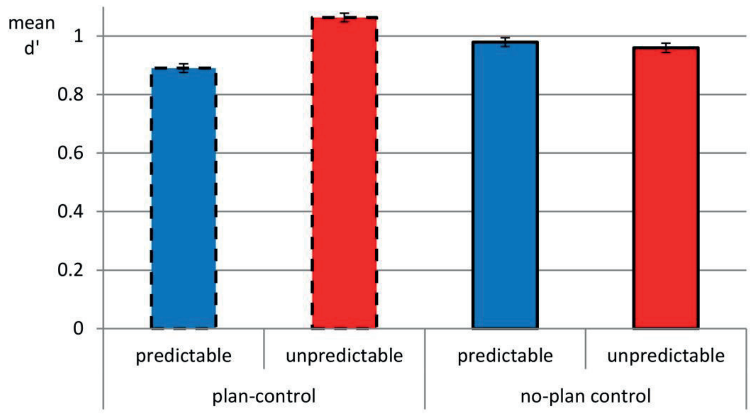

Figure 6. Mean d' value by task (plan-control, no-plan-control) and predictability in the recognition memory test of Experiment 2. Error bars reflect the standard error of the mean.

Decision speed was evaluated with the same model as in Experiment 1. Outliers were set to two standard deviations from the participant mean (4.76 % of the data excluded). The analysis revealed only a main effect of response (i.e., “new” or “old” responses; note that again these correspond to misses and hits, as only targets and not lures were analysed; b = -97.13, p <.0001, intercept = 962.21) and not of task (b = 12.62, p =.102) or predictability (b = 19.17, p =.092). "Old" responses (844 ms) were given faster than "new" responses (1078ms). In addition, there was a significant two-way interaction between predictability and response (b = 19.54, p =.013). We ran two additional models; one for the level "old" and for the level "new" of the factor "response". Predictability (predictable vs unpredictable sentence-final-word) was important when labeling an item as "old" (meaning, when making a hit; b = 36.43, p =.001, intercept = 872.66), whereas when labeling it "new" (i.e., a miss) there was no such effect (b =.40, p =.981, intercept = 1085.40). In particular, unpredictable last words were labeled "old" significantly faster than predictable ones (810 vs. 881 ms; see also Figure 7.

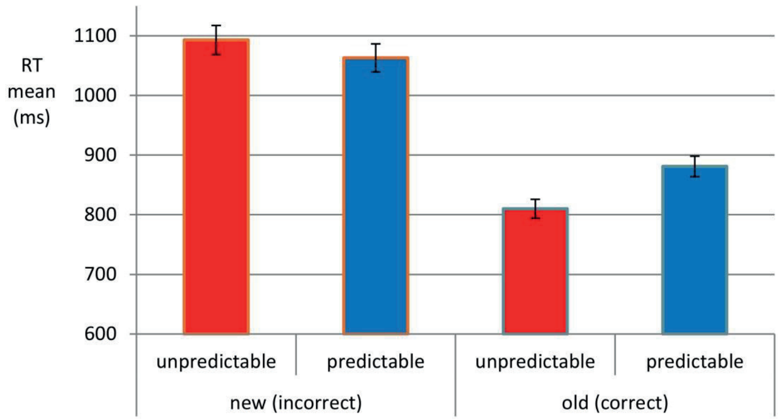

Figure 7. Mean decision speed (ms) for the recognition memory test of Experiment 2 as a function of response and predictability. Error bars reflect the standard error of the mean.

Pupil Size Data. For the analysis of mean and peak pupil size (again only targets and not lures were analysed), the fixed structure of the tested models included response ("new" and "old", fitted as -1 and 1 respectively), task (plan-control and no-plan-control, fitted as -1 and 1 respectively), predictability (unpredictable and predictable, fitted as -1 and 1 respectively), plausibility (centered), experimental list (centered) and the interaction of response, task and predictability. Random intercepts by subjects and by spoken words and a random slope for task and predictability by subjects and by spoken words made up the random structure of the model. These tests revealed only a main effect of response (mean pupil size model: $b = .006$, $p < .0001$, intercept $= .041$, peak pupil size model: $b = .008$, $p < .0001$, intercept $= .097$), with "old" responses (meaning hits) leading to larger dilation than "new" responses (meaning misses).

Comparison of Experiment 1 and 2

We compared the two experiments in only two of the dependent measures; Mean and peak pupil size in the exposure phase and d prime values in the recognition memory test. The first comparison was informative on whether the pupil size differences we observed in Experiment 1 were indeed the result of planning effort. The second comparison verified whether the recognition memory decrement in Experiment 1 was truly the result of concurrent speech planning while listening. We report only data that concern main effects of experiment or the interaction of experiment with another factor.

***Exposure phase*.**

Pupil Size Data. The initial model for mean and peak pupil size included Experiment (Experiment 1 and Experiment 2, fitted as -1 and 1 respectively), task (plan, no-plan, fitted as -1 and 1 respectively; note that for Experiment 2 the task labels are "plan-control" and "no-plan-control"), predictability (unpredictable and predictable, fitted as -1 and 1 respectively), plausibility(centered), experimental list (centered), as well as the three-way interaction of Experiment, task and predictability and their two-way interactions as fixed factors. The random structure included random intercepts by subjects and picture-names, and a random slope for task and predictability and their interaction by subjects only. The final model indicated that Experiment (mean pupil size model: $b = .029$, $p < .0001$, intercept $= .012$, peak pupil size model: $b = .038$, $p < .0001$, intercept $= .057$) contributed significantly to model fit, as pupil dilation was larger for experiment 1 than experiment 2. In addition, an interaction of Experiment with Task (mean pupil size model: $b = -.040$, $p < .0001$, peak pupil size model: $b = -.046$, $p < .0001$) was found. The separate analyses reported in the experiment-specific sections indicate that this interaction results from the fact that task modulated pupil size in Experiment 1 (Figure 1), but not in Experiment 2 (Figure 4).

Recognition memory test. The sensitivity measure (d prime) was submitted to two-way mixed Analyses of Variance (ANOVA) by participants. Task (plan-control, no-plan-control) and predictability (predictable, unpredictable) were entered as a within-participant factor and experiment as a between participants factor. The analysis revealed a significant effect of Experiment (F_1 (1, 62) = 9.368, p =.003, η^2 =.131), as overall memory was better for experiment 2 compared to experiment 1, and of the interaction of Experiment with task (F_1 (1, 62) = 23.895, p <.001, η^2 =.278). Based on the experiment-specific analyses reported above, this interaction results from the fact that there was a difference in d prime magnitude between plan and no-plan items of Experiment 1, but not in Experiment 2 (see also Figures 2 and 6).

General Discussion

The first experiment of the current study investigated whether planning to speak negatively affects the processing of a concurrent spoken word in sentence final position. As such, the experiment extends the design of chapter 3 (which focused on the processing of single words) to a more natural setting. The experiment also manipulated the predictability of the sentence-final-word to investigate whether the potential detrimental effects of concurrent naming on listening (as evaluated via recognition memory performance) could be (partly) alleviated by this factor. This is important because in natural conversation sentence-final-words may vary in their predictability. Finally, a control experiment (Experiment 2) tested whether the observed effects could have been caused by a difference in stimulus variability between the plan and no-plan conditions (i.e., independent of the planning process itself), thereby controlling for a potential confound in the design. The results of this experiment indicate that the effects observed in Experiment 1 are indeed the product of interference induced by speech planning.

Interference between speaking and listening

The results of this study demonstrate that preparing to name a picture has a substantial impact on recognition memory for heard words. This is evident from the findings in Experiment 1, where participants' recognition memory was worse for items that were heard while participants were planning to name a picture compared to items that participants heard when they were not engaged in planning. This observation was corroborated by the fact that it took participants longer to correctly recognize an item as "old" when that item had initially been heard while they were planning, compared to deciding on items initially heard in the no-plan task.

The second experiment demonstrated that these findings could not be attributed to a mere difference in novelty/variability between pictures in the plan versus the no-plan tasks. In the first experiment a strong effect of speech planning on listening was observed as a decrement in recognition memory for items that had been heard during speech planning. In Experiment 2, where participants were instructed to ignore the pictures, we found no difference between words heard in conjunction with nameable pictures (i.e., the pictures from the plan task in Experiment 1) and those heard in conjunction with the meaningless line drawing (i.e., the visual stimulus shown during the no-plan task in Experiment 1). The difference in findings between Experiment 1 and 2 shows that the memory decrement we observed in Experiment 1 for the plan task was indeed driven by planning effort and not by other factors like picture recognition per se, or novelty of the visual stimulus (picture) as opposed to a constant visual stimulus (meaningless line drawing).

This conclusion is further supported by the pupil dilation data - a known index of processing load (e.g. Beatty & Lucero-Wagoner, 2000; Kahneman, 1973; Laeng, Sirois, &

Gredeback, 2012)- in the exposure phase. In general, we observed reliable modulations of pupil size as a result of planning effort. For example, Experiment 1 demonstrated that pupil dilation was modulated by picture-name frequency, with low frequency picture-names leading to larger pupil dilation. More importantly, pupil size was larger for plan compared to no-plan trials during the exposure phase of Experiment 1. No such modulation was observed in Experiment 2, for which plan-control and no-plan-control tasks did not differ in pupil dilation, as did not high- and low- frequency picture-names. This demonstrates that picture naming was effortful; a finding that more directly relates the processing costs of planning in the exposure phase to the decrement in memory performance at test.

Although people naturally switch between speaking and listening numerous times in typical conversational settings, the findings presented here thus suggest that speaking and listening are tasks that compete for cognitive resources. This notion is supported by our recognition memory results in Experiment 1, but also by the comparison between Experiments 1 and 2. Participants performed better in the recognition memory task of Experiment 2 (where they were instructed to listen to the sentences and only passively view the pictures that appeared on screen), than when instructed to listen and name the pictures (Experiment 1). In addition, pupil dilation was larger in Experiment 1 than in Experiment 2 and task modulated pupil size only in Experiment 1. In addition, and although subtle, the observation that predictability affected naming latencies, also demonstrates that our manipulation of perception-difficultly (perceiving a predictable as opposed to an unpredictable sentence-final-word) did have cognitively relevant effects on production. These findings further strengthen the conclusion that planning interferes with listening.

CHAPTER 4: SPEECH PLANNING AFFECTS MEMORY FOR SENTENCE FINAL WORDS

Predictability of sentence final words

For Experiment 1, in addition to the manipulation of concurrent planning, we had hypothesized that when preparing to name a picture, a sentence-final-word that could be predicted from prior context, would be remembered better than one that could not be predicted from prior context. The idea was that despite planning, being able to predict the sentence-final-word from the preceding context would alleviate the cost in processing resources for listening. We did not find support for this hypothesis. Items that were predictable were equally well remembered as the unpredictable items in both the plan and the no-plan tasks. Meyer, Mecklinger, and Friederici (2007) also report no difference in accuracy or decision speed in a recognition memory test for correct versus semantically violating verbs, while the N400 data suggested that indeed these had been processed differently. Maybe using an indirect memory measure, like repetition priming magnitude, rather than a direct memory measure, like recognition memory performance, would have picked up a difference between predictable and unpredictable items (see Bentin, Kutas, & Hillyard, 1995 & Schacter, 1987 for this discussion).

In contrast to this lack of an effect of predictability on memory performance, an effect of predictability was observed in the naming latencies during exposure. Participants were faster to name low frequency pictures when the sentence-final-word was predictable. No effect of predictability was seen for high frequency pictures. Thus facilitation by predictability on concurrent naming is only captured when naming is more effortful, as is the case for low frequency picture names. Listening to a predictable sentence-final-word then indeed seems to be less effortful (than listening to an unpredictable sentence-final-word), thereby leaving more available resources for preparing to name a demanding low frequency picture; or at least allowing for a faster switch of attention from listening towards planning. Listening to an unpredictable sentence-final-word might thus be more effortful (than listening to a predictable sentence-final-word), thereby leaving less available resources for preparing to name a demanding low frequency picture; or at least not allowing for a fast switch of attention from listening towards planning, since the sentence-final-word first needs to be perceived. Thus, our predictability manipulation might not have affected recognition memory, but it did impact on planning speed, when planning was most effortful.

A related observation is that in Experiment 2, where no naming was involved, it was shown that unpredictable words were remembered more often than predictable ones on trials with nameable pictures (i.e., the pictures that featured on the plan trials in Experiment 1; although participants were instructed to ignore these for Experiment 2). Moreover, unpredictable sentence-final-words were labeled "old" significantly faster than predictable ones during the recognition memory test. At least two studies (Corley, MacGregor, & Donaldson, 2007; MacGregor, Corley, & Donaldson, 2010) report a similar effect, with unpredictable utterance endings being more likely to be remembered than predictable ones. Two potential explanations could account for this observation.

A first explanation has to do with the notion of interference. According to this explanation prediction makes listening easier. As a result, more capacity may be left available to process the concurrent pictures (i.e., even though participants were instructed to ignore these). That is, participants might occasionally have engaged in more elaborate processing of the pictures (maybe even covert naming), thereby inducing some interference with the heard speech. As a result memory performance dropped. When participants could not predict, then that capacity would have been directed to listening, making it easier to ignore the picture that appeared on screen and avoid interference. As a result they could have better memory of the sentence-final word in the unpredictable situation. This explanation also matches our observations on how naming high and low frequency picture-names is affected by predictability.

Alternatively, the better memory for unpredictable sentence-final-words might not link to interference but merely be the product of an element of surprise that comes with unpredictable items, which might have prompted participants to just listen more thoroughly compared to the predictable items and thus also better encode these unpredictable items in memory. We deem this second explanation less likely, however, because: 1) these sentence-final-words although unpredictable were not really surprising but rather still plausible; and 2) more thorough listening due to unpredictable items also predicts that pupil dilation would differ between predictable and unpredictable items, which was not the case (see Tromp, Hagoort, & Meyer, 2016 for a demonstration on how pupil dilation is affected by listening effort). Note that following the first explanation pupil dilation does not need to differ between predictable and unpredictable items, since the resources freed by listening to a predictable word might have been immediately invested into processing the picture and maybe even covert naming, thereby leading to comparable pupil sizes between predictable and unpredictable items. Clearly, more research is needed to locate the actual source of this unanticipated effect.

The source of the interference.

Preparing to speak takes away resources from listening, and, under some circumstances, effortful listening (unpredictable words) takes away resources from planning your speech. Although the findings presented here demonstrate interference effects between speech planning and speech perception, the exact locus of this interference remains unclear.

In our study we adopted the concept of general cognitive resource sharing that has been used as an explanation for the so-called Psychological Refractory Period (PRP) phenomenon (e.g., Sommer & Hohlfeld, 2008). PRP refers to the observation that when two stimuli (S1 and S2), to which a reaction is needed, are presented very close in time, the reaction time to the second stimulus (S2) is dependent on the stimulus onset asynchrony (SOA, typically ranging between 50 ms and 1 s) between the first (S1) and second stimulus (S2), suggesting that processing of the second stimulus is either postponed (bottle-neck accounts: e.g. Pashler, 1984, 1994) or slowed down (resource sharing accounts: e.g. Navon & Miller, 2002; Tombu & Jolicœur, 2005).

The design of our study did not strictly follow the PRP paradigm and we are not in position to differentiate between the two accounts. As such, our choice of the term "resource sharing" does not imply that we embrace the resource sharing accounts as opposed to the bottle-neck accounts as an explanation for the observed interference. Our study did however involve a dual tasking situation in which the SOA between S1 (sentence-final-word) and S2 (picture) was zero. Thus, the observed decrements in recognition memory due to the interference between tasks may be linked to the PRP phenomenon. Note however that in a PRP framework one would focus on how S2 is affected by S1, while in our study we mainly focus on the opposite; namely on how S1 (listening to the sentence-final-word) is affected by S2 (planning the picture-name). Yet an impact of S1 on S2 is still visible in our study, as evidenced by the impact of predictability of S1 (sentence-final-word) on planning speed of S2 (low frequency picture names; Experiment 1). As such, our finding that low frequency picture-names were named faster when the sentence-final-word was predictable, might be linked to such a central processing bottleneck. If indeed listening to the predictable sentence-final-word was easier than listening to the unpredictable one, then processes for S1 predictable (predictable sentence-final-word) ended faster than processes for S1 unpredictable (unpredictable sentence-final-word) and as a result S2 was processed faster when S1 was predictable than when it was unpredictable.

The psycholinguistic literature on dual-task interference has at times adopted the bottle-neck account and at times the resource sharing account. For example, research that has looked into how producing a picture-name affects performance on a concurrent tone discrimination task has made use of the single channel bottleneck account and has argued that lemma and phonological word-form selection are subject to a central processing bottleneck, whereas phoneme selection is not (Ferreira & Pashler, 2002). Similarly Hohlfeld, Sangals, and Sommer (2004) argued in favor of single channel processing models, based on their findings that the N400 component is delayed, when participants have to cope with a dual task situation. That is semantic processing is postponed by additional tasks. An interesting finding of this study was that language-related stimuli (letters) led to an even more pronounced delay compared to spatial stimuli. We return to this point further down.

When Hohlfeld and Sommer (2005) observed a reduction rather than a delay in the N400 they concluded that interpreting a reduction in amplitude would rather implicate some kind of resource sharing, whereby attentional resources are withdrawn from the semantic properties of words, which are only attended to with respect to the characteristics needed by the task (Sommer & Hohlfeld, 2008; see also Hohlfeld & Sommer, 2015; Luck, 1998). It seems thus that both accounts can contribute to interpreting performance in psycholinguistic tasks.

Going back to the study by Hohlfeld and colleagues (2004), an interesting finding was that task interference reached its maximum when the overlapping tasks both involved processing language-related items. This seems to suggest that apart from the difficulty of having to share general cognitive resources between the tasks per se, the nature of the overlapping representations might also add to the equation. It has been argued that the neuronal infrastructures of speaking and listening show substantial overlap at the lexical level (Menenti, Gierhan, Segaert, and Hagoort, 2011). If this is indeed the case, then in the current study interference could get introduced not only due to the need to share general cognitive resources but also due to overlapping linguistic representations. Lexical retrieval for naming a picture may therefore interfere with lexical retrieval processes for listening to the constituents of the sentence. This overlap might add to the interference, which could in turn lead to memory deterioration. Note however that it cannot only be lexical interference of overlapping representations that produces the memory decrement. In our study we observed that predictability differentially affected low frequency pictures compared to high frequency pictures. Such an effect cannot be solely attributed to lexical interference due to overlapping representations, since both low and high frequency picture-names could suffer from lexical interference.

Finally, a question remains whether actual perception of the heard word or only the storage in memory is affected by concurrent planning. Although this issue cannot be fully resolved with the current findings, our observation that recognition memory performance drops when attention is divided during encoding matches the literature on divided attention and memory performance. According to this literature dividing attention during encoding leads to poorer recognition- (e.g. Fernandes and Moscovitch, 2000) and recollection-memory performance (e.g. Naveh-Benjamin, Guez, & Sorek, 2007). It seems that dividing attention between listening and preparing to speak leads to a shallower encoding of the sentence-final-word. Naveh-benjamin, Craik, Gavrilescu, and Anderson (2000) have proposed that dividing attention during encoding affects the quality of encoding, changing it from a semantically elaborate type of processing to a shallower type. As they conclude, poorer memory performance due to divided attention is the result of a quantitative shift in encoding, towards less deep elaborative strategies.

Conclusion

Preparing to name a picture while still listening to the final word of a sentence comes with a cost for memory encoding of the sentence-final-word. Recognition memory was shown to be poorer for items that were encoded during planning, compared to items encoded during listening only. This loss in memory performance seems to be linked to elevated cognitive load during planning while listening, as evidenced by changes in pupil dilation. The findings of the current study suggest that concurrent listening and speech-planning is not as easy as it might seem in conversation. Interlocutors have to learn to cope with the cognitive load induced by this parallel processing situation. They have to be able to effectively distribute cognitive resources between tasks, in order to reach maximum listening quality (at least as evaluated via recognition memory performance) while also keeping the smooth pace of the conversation.

Footnotes

[1] The random structure is described once for each dependent measure. Only if convergence problems forced us to alter it in any way it is mentioned again. When not explicitly stated, the random structure described at the beginning of each section applies to all the models of this dependent measure.

[2] The estimates correspond to ratios as measured from relative pupil size and not to percentage of pupil change.

[3] We use d' rather than other measures as a measure of accuracy, because d' is a robust measure that depends on stimulus parameters and remains roughly invariant when response bias is manipulated. This is not the case for measures like proportion correct.

Chapter 5

Speech planning affects offline and online processing of spoken words

Gerakaki, S., Sjerps, M.J., Rommers, J. Speech planning affects offline and online processing of spoken words.

Abstract

In conversation, interlocutors are often planning their speech while listening to each other. In chapters 3 and 4 we demonstrated that concurrent speech planning negatively impacts listening quality. Yet, with the exception of pupil dilation, the measures used (reaction times and recognition memory performance) monitored performance offline, rather than online. The current chapter investigated whether and how the quality of listening is affected both offline and online by concurrent speech planning. In an initial part of the experiment (exposure phase) participants heard sentences that ended in an expected ("With tea we always eat a <u>cookie</u>") or a semantically anomalous ("With tea we always eat a <u>mouse</u>") sentence-final-word. On half the trials, a picture was presented at sentence-final-word onset, which participants had to name (plan task). On the remaining trials they saw a meaningless line drawing instead of the picture and remained silent (no-plan task). EEG was recorded, as well as picture naming latencies. A second phase (recognition-memory test), evaluated how well participants remembered the sentence-final-words of the exposure phase.

As expected, during the no-plan task, the N400 was reduced in response to expected words compared to anomalous words. Importantly, online (EEG) and offline (recognition memory) measures indicated that listening was affected by the concurrent task: ERPs that were elicited during the sentence-final-words of the exposure phase revealed a substantial decrease in the amplitude of the N400 semantic anomaly-effect for items heard during speech

planning compared to items heard when not engaged in planning. Moreover, recognition memory was poorer for sentence-final-words heard during the plan than in the no-plan task. These results demonstrate that during speech planning, the encoding of heard speech is impaired as it does not properly activate semantic representations, and thus lacks the formation of an easily retrievable memory trace. Performance in the recognition memory test further showed that not only planning affected the quality of sentence-final-word-processing but also the congruency manipulation of the sentence-final-words. Among the subset of sentence-final-words that had been presented on picture naming trials, those words that were expected given their preceding context were better remembered than the semantically anomalous ones. Thus, listeners may alleviate some of the cost induced by planning, by relying more heavily on the sentence context in which words had been presented.

CHAPTER 5: SPEECH PLANNING AFFECTS PROCESSING OF SPOKEN WORDS

Introduction

When engaged in conversation, interlocutors switch between speaker and listener role. It has been demonstrated that at transitions of turns interlocutors often start planning their turn before the turn-end of their interlocutor (Bögels, Casillas, & Levinson, 2018; Bögels, Magyari, & Levinson, 2015; Boiteau, Malone, Peters, & Almor, 2014; Sjerps & Meyer, 2015). This raises the question to what extent concurrent listening and speaking may result in interference.

Chapters 3 and 4 in this thesis demonstrated that concurrent planning and listening can indeed result in cognitive interference between listening and planning. In these chapters listening quality was mainly monitored offline with reaction time- and recognition memory performance-measures, while only pupil dilation measures offered an online monitoring tool for the cognitive effort involved in planning. Yet no measure offered an insight on online processing during listening, a measure that would demonstrate that the impact of planning on listening arises in real time. Therefore this chapter focused on providing information on how online listening quality is affected by concurrent speech planning, while still evaluating listening quality offline. Online listening performance was evaluated with electroencephalography (EEG) recordings and offline with behavioral measures.

Furthermore, in typical speech settings, speech planning especially overlaps with sentence-final-words, which are often predictable from the prior context. In chapter 4 we had hypothesized that predictability of the sentence-final-words could alleviate part of the cost induced by concurrent planning. Possibly the type of predictability manipulation we introduced in chapter 4 was too subtle to allow for such effects to be observed. Therefore, the current chapter implemented a more extreme predictability manipulation by contrasting expected and anomalous sentence-final-words, rather than predictable versus unpredictable sentence-final-words (chapter 4). Such a predictability manipulation might indeed interact with, or even alleviate, the potential detrimental effects of concurrent speech planning.

Performance loss and processing tradeoffs in dual tasking.

Speech production and speech perception processes interact at multiple levels of processing. For example a number of studies using the picture-word-interference task have shown that the presentation of written or spoken words can affect single word production due to semantic or phonological relatedness (Damian & Martin, 1999; Schriefers et al., 1990). In addition to such representation-based interactions, however, speech perception and production processes may also interact as a result of the limited set of domain-general cognitive resources that have to be shared between any two or more tasks (e.g. Baddeley, 1976; Becic et al., 2010; Kemper, Herman, & Lian, 2003; Kemper, Schmalzried, Herman, & Mohankumar, 2011; Lavie, Hirst, Fockert, & Viding, 2004; Lavie, 2005; Meyer & Kieras, 1997; Pashler, 1984, 1994). That is, when a limited amount of cognitive capacity is distributed across multiple tasks, performance in the two tasks typically involves a direct tradeoff. Increases in performance on one task are correlated with decreases in the secondary task (Somberg & Salthouse, 1982). As argued in the chapters 3 and 4 of this thesis, such factors are also likely to play a role in situations of concurrent listening and speaking.

The consequences of having to share general processing resources between linguistic and nonlinguistic tasks have been demonstrated to affect language related processes at multiple levels in research investigating perception and production tasks separately but in combination with nonlinguistic tasks. For example, in a combined picture naming and concurrent secondary tone-discrimination task, performance on both tasks is modulated by the ease of the picture lemma or word form selection (Ferreira & Pashler, 2002). Moreover, influences of central capacity demands on speech production have been observed for phonological processing (Cook & Meyer, 2008; see Roelofs & Piai, 2011, for a review on attention and spoken word planning). Similarly, effects of dual task interference with nonlinguistic tasks have also been observed across different levels of perception. Secondary nonlinguistic tasks affect the perception of individual phonemes (e.g. Gordon, Eberhardt, & Rueckl, 1993; Mattys & Wiget, 2011), individual words (e.g. Cleland, Tamminen, Quinlan, & Gaskell, 2012), and the perception of sentences (Bosker et al., 2017).

And beyond these relatively local effects on single word perception and production tasks, interference effects on more naturalistic language behavior has also been demonstrated. For example, language production and perception tasks have a strong impact on concurrent driving (e.g., Kubose et al., 2006; see also Strayer & Johnston, 2001), and the reverse is true as well, as driving reduces concurrent language production and comprehension performance (Becic et al., 2010). That is, Becic et al., observed poorer story retelling performance, and a negative impact on story comprehension and long-term memory encoding, when conversing while driving. These findings demonstrate that both speech production and perception interfere with secondary nonlinguistic tasks. Hence, it seems likely that speech perception and production may also interfere with each other in such a domain-general sense.

Assessing online listening quality: The N400 in adverse conditions.

The current chapter aimed to investigate to what extent the processing of heard words is affected by concurrent speech production planning, with a specific focus on its online effect on semantic processing. We measured processing quality online by eliciting Event Related Potentials (ERPs). Our ERP analyses focused on the N400 component (Kutas & Hillyard, 1980) which is considered to be the electrophysiological signature of semantic processing. It is a negativity arising between 200 and 600 ms, with a peak around 400 ms and a centroparietal distribution with a slight right-hemisphere bias (for visually presented language; Kutas & Federmeier, 2011). The N400 is elicited by any potentially meaningful stimulus, but its amplitude is reduced to the extent that there is contextual support for the stimulus, suggesting that it reflects degree of contextual facilitation (Delong, Urbach, Groppe, & Kutas, 2011). We asked whether the auditory input will still be processed at a semantic level (triggering an N400), when the sentence-final-word is not the sole focus of attention.

Several studies have focused on how divided attention reduces -or even eliminates- N400 effects (see Kutas & Federmeier, 2011; and Van Petten, 2014, for review). For example, presenting words in an unattended spatial location, or in a font color that participants are told to ignore, results in a reduced or even completely absent N400 (Bentin et al., 1995; Kellenbach & Michie, 1996; Mccarthy & Nobre, 1993; Vogel, Woodman, & Luck, 2005). In many of these studies, a reduced N400 was observed along with reduced subsequent recognition memory, suggesting a close relationship between the N400 and the formation of easily retrievable memory traces (see also Bentin, Kutas, & Hillyard, 1995; Heil & Rolke, 2004; Kellenbach & Michie, 1996; Otten, Rugg, & Doyle, 1993; Phillips & Lesperance, 2003). It is important to note, however, that although an absence of the N400 is related to a reduction in recognition memory, it does not imply a complete absence thereof. For example, Kellenbach & Michie, 1996 report that despite the absence of an N400, recognition memory was still better than chance (see also Heil & Rolke, 2004; Otten, Rugg, & Doyle, 1993; Phillips & Lesperance, 2003). Together, however, and most relevant to the current experiment, these findings suggest that the N400 is sensitive to attentional modulations.

In addition, the N400 priming effect has been found to depend on task demands. A lexical decision task that included a relatedness manipulation induced an N400 response, while a task including the same relatedness manipulation, but in which subjects had to discriminate between lowercase and uppercase letters, did not (Chwilla, Brown, & Hagoort, 1995). Still, reliable N400 priming effects were observed when the task required phonological or orthographic processing instead (Connolly, Stewart, & Phillips, 1990) and when the task required grammatical discrimination or a letter search (Küper & Heil, 2009). It seems thus that semantic processing is rather robust and remains unaffected by a number of task manipulations. Presumably not all tasks create such high cognitive workload that could impair semantic processing as reflected in the N400 effect.

Attending to multiple tasks at once and thus possibly having to share cognitive resources between tasks, could be a way to induce such high cognitive workload to impair semantic processing. For example, Luck, Vogel, and Shapiro (1996) had participants complete on each trial both a decision on whether the presented number was odd or even and on whether the presented word was related or unrelated to the context word of that trial. They used rapid serial visual presentation (RSVP) to induce the attentional blink effect. This is the observation that participants often fail to report the second in line target within a stream of distractors, if this is presented within 200 to 500 ms of the first target (Dux & Marois, 2009). Their study revealed that even though accuracy in detecting semantic relatedness was subject to the attentional blink, the N400 amplitude was not (see also Rolke, Heil, Streb, & Hennighausen, 2001 for a similar pattern when the context word is presented during the attentional blink).

A number of more recent studies using a comparable experimental paradigm to Luck and colleagues (1996) report a delayed or attenuated N400 (see Batterink, Karns, Yamada, & Neville, 2010; Giesbrecht, Sy, & Elliott, 2007; Lien, Ruthruff, Cornett, Goodin, & Allen, 2008). Two of these recent studies make an interesting assumption that semantic processing is affected by task switching during which the task set needs to be reconfigured (Vachon & Jolicoeur, 2011; Vachon & Jolicœur, 2012). When this has to happen fast (due to short SOA) the effect is even stronger. Even switching from one semantic task (object classification judgment) to another semantic task (semantic relation judgment) reduced N400 amplitude (Vachon & Jolicœur, 2012). On a similar note, Hohlfeld, Sangals, and Sommer (2004) reported a delay of the N400 and concluded that the completion and possibly even the initiation of semantic processing had to be postponed because participants had to first complete some other task before turning to the semantic relatedness task. The delay was even more pronounced when the other task required responding to language-related stimuli (letters) compared to spatial stimuli.

In the case of planning speech while listening to speech reconfiguring the task set might involve switching attention away from listening and towards speaking preparation. As such, planning speech while listening to speech might be seen as a dual tasking situation of high temporal overlap that could be expected to result in performance loss for one or both modalities. At the same time it is a far more practiced task than the tasks in the up to now cited studies. Importantly, most of the existing studies focused on single-word level relatedness judgments. Yet task demands might have more impact on semantic processing when the task involves processing single words rather than whole sentences (see also Kutas & Federmeier, 2011). Therefore, it is not obvious how production planning might affect listening and semantic processing in a sentential context.

Memory performance and its interaction with task manipulations and semantic congruity.

Apart from evaluating online semantic processing by using the N400, we were also interested in an offline measure of listening quality. To this end we investigated the strength of memory formation of the heard words offline, using a recognition memory task, as in chapters 3 and 4. This was done because previous research has shown that when attentional manipulations are involved, semantic processing -as witnessed by tracing N400 activity- may endure, while the processed items might not reach the level of a durable memory in a later recognition memory test (Kellenbach & Michie, 1996; Olichney et al., 2000; Phillips & Lesperance, 2003). Thus both on- and offline measures are important in establishing performance loss on the listening side. In general, dividing attention during encoding has been shown to lead to poorer recognition- (e.g. Fernandes and Moscovitch, 2000) and recollection-memory performance (e.g. Naveh-Benjamin, Guez, & Sorek, 2007; and see other chapters in this thesis). As reviewed in the previous section, research that has investigated how the N400 is affected by task manipulations has on occasion also examined memory performance. Such studies have revealed that attended items are generally remembered better than unattended or ignored items in a subsequent recognition memory task (Bentin, Kutas, & Hillyard, 1995 and Kellenbach & Michie, 1996).

Not only task manipulations but also semantic congruity interacts with memory performance for words. For example, some studies report better memory for semantically congruous than incongruous words in both a recall- and a recognition- memory test (Olichney et al., 2000; see also Neville, Kutas, Chesney, & Schmidt, 1986). Superior recall memory performance for congruous compared to incongruous sentence-final-words has also been reported by Besson, Kutas, & Van Petten (1992). Thus, task manipulations of this kind show that semantically related items tend to be recalled better than unrelated.

In contrast to these studies, Corley, MacGregor, and Donaldson (2007) found that unpredictable words were remembered better than predictable words. To look at how hesitations affect comprehension they had participants simply listen to conversation extracts that were interspersed with hesitations. After completing the comprehension task they took part in a recognition memory task, in which they heard the utterance-final target words of the previous task intermixed with lures. Items that were preceded by a hesitation were remembered better, as were unpredictable items overall. The same is reported in a subsequent study by MacGregor, Corley, and Donaldson (2010). Both studies used highly constrained sentences ending in predictable or unpredictable (cloze probability = 0) target words (see table 1, page 661 of Corley et al., 2007 for example sentences). An advantage in recognition memory performance for unexpected items is also reported by Wlotko, Federmeier, and Kutas (2012) who compared strongly- and weakly- constraining sentence frames with expected and unexpected (but still plausible) sentence endings.

Thus, it seems that depending on the task manipulation and the type of stimuli used, either expected or unexpected items lead to superior memory performance. In chapter 4 we had hypothesized that participants would demonstrate superior recognition memory performance for predictable items, compared to unpredictable, when participants were engaged in planning while listening. One reason why this prediction was not borne out might relate to the type of stimuli used. These involved constraining (e.g. "Het schip werd stevig vastgelegd met een **anker**" /"*The ship is firmly established with an **anchor***") and non-constraining sentence-frames (" Dat zware metalen stuk daar is een **anker**" /"*This heavy metal thing there is an **anchor***"). As such these sentence-final-words were either highly expected (constraining sentence-frame) or rather unexpected (non-constraining sentence-frame), but still plausible. Perhaps this congruity manipulation was too weak to interact with task manipulation (plan versus no-plan in chapter 4). Perhaps a more extreme congruity manipulation contrasting an expected to a semantically anomalous sentence-final-word could pick up such an interaction. As a result in the current chapter we investigated again the role of predictability in alleviating potential memory performance costs due to concurrent planning. Only this time we compared expected to semantically anomalous words, which can be considered a more extreme congruity manipulation compared to chapter 4.

Summary.

In sum, in both the literature on dual tasking and the literature on the N400, memory and attentional manipulations seem to suggest that under dual tasking conditions, where one has to start planning while still listening, there might be some performance loss on the planning or the listening side. The studies reported in chapters 3 and 4 of the thesis support this view. We aim to investigate whether indeed performance loss can be traced to the listening side both off line (behavioral measures) and online (EEG) and whether making listening easier by providing an expected rather than an anomalous sentence-final-word will alleviate this performance loss.

Paradigm and hypotheses.

The experiment consisted of two phases; an exposure phase and a recognition memory test. In the exposure phase participants listened to Dutch sentences that ended either in a semantically expected (for example "With tea we always eat a <u>cookie</u>") or a semantically anomalous ("With tea we always eat a <u>mouse</u>") last word (hereafter sentence-final-word). On plan trials, a picture appeared on screen at the onset of the last word. In trials to which participants did not have to respond (no-plan trials) the picture was replaced by a meaningless line drawing (see Figure 1). Participants had to name the pictures as quickly as possible, but they also had to listen carefully to the sentences. They were told that they would later "have to do something with these sentences". Participants were not specifically instructed to pay attention to the last word of the sentence (i.e., the critical words).

Response time to the pictures was evaluated as a function of the predictability of the auditorily presented words. We hypothesized that pictures that appeared with an expected sentence-final-word would be named faster than those with an anomalous sentence-final-word, since either the expected words should be easier to process or the anomalous sentence-final-words might take more listening effort.

To assess listening performance online, we investigated semantic processing by recording EEG. Modulation of the N400 amplitude was evaluated as a function of task (plan, no-plan) and predictability. If preparing to name a picture while still listening to a sentence is comparable to the dual task situations experimentally induced in studies like those of Hohlfeld and Sommer (2005) and Vachon and Jolicoeur (2011, 2012), we expected to find an interaction effect, by which the semantic anomaly effect (the difference between anomalous and expected sentence-final-word) would become smaller in the plan compared to the no-plan task, because participants were preparing speech while listening to the sentence-final-word.

After the exposure phase, participants completed a recognition memory task. The same paradigm was used as described in chapter 4. Participants heard all last words of the sentences which they had heard during the exposure phase intermixed with lures and had to indicate whether or not they had heard each word in the exposure phase. We evaluated memory performance (d prime) and decision speed as a function of task (plan, no-plan) and predictability. In accordance with studies showing that unattended items are remembered less often than attended items (Bentin et al., 1995; Kellenbach & Michie, 1996), we expected that sentence-final-words heard during planning speech should be remembered less often than those heard in a no-plan setting.

Given that the studies on memory performance cited above have produced mixed results (Besson et al., 1992; Corley et al., 2007; Kellenbach & Michie, 1996; MacGregor et al., 2010; Olichney et al., 2000; Phillips & Lesperance, 2003; Wlotko et al., 2012) with superior memory being reported both for predictable and for unpredictable target items -and since our research questions were focused on the interaction with planning-, we did not make a specific prediction on the direction of the effect, but we did predict a main effect of predictability for the recognition memory test. What we did expect to find was an interaction of task (plan, no-plan) with predictability, whereby sentence-final-words that were expected should be less affected by having to plan speech, than those that were anomalous. As already argued in the introduction section, we expect to find such an effect, despite failing to observe it in chapter 4, because we assume that in an environment where planning takes away resources from listening, remembering an anomalous sentence-final-word that does not match the sentence context should be more difficult than remembering an expected sentence-final-word that comes as a confirmation of what the sentence context already points at.

For decision speed the hypotheses paralleled memory performance. The focus here was on responses in which participants correctly label an item "old". Participants should be faster in deciding that an item is old, when this was presented in the no-plan compared to the plan task and it was an expected, compared to anomalous sentence-final-word.

Method

Participants. 38 participants were recruited from the participant pool of the Max Planck Institute for Psycholinguistics, Nijmegen. They received a monetary reward for their participation. Fourteen participants had to be excluded due to fatigue (n = 4), technical difficulties (n = 2), or applying a too conservative or too liberal decision criterion during the recognition memory task (n = 8; participants with lambda values below 0 or above 1.5 were excluded; Wickens, 2001). All participants were adult native speakers of Dutch (mean age = 22.8 years, 17 female), had normal or corrected-to-normal vision, no known speech or hearing problems, and were not diagnosed with dyslexia. All study protocols were approved by the Ethics Board of the Social Sciences Faculty of the Radboud University Nijmegen.

Materials and Design. For the exposure phase materials, 164 sentences were selected from those used by Piai, Roelofs, and Maris (2014, 61 sentences) and by Rommers, Meyer, Praamstra, and Huettig (2013, 99 sentences). Four of those were used as practice material. 160 were used for the experimental exposure phase task. Two versions of each sentence were generated; ones that ended in an expected final word (e.g., "Bij de thee eten we altijd een koekje " [With tea we always eat a cookie]) and ones that ended in a semantically anomalous word (e.g., " Bij de thee eten we altijd een muisje" [With tea we always eat a mouse]; see appendix J for a listing of the sentences). For the expected final words we did not collect cloze probability measures but they were a subset of the sentences for which the average cloze probability reported was.72 (SD =.30) in Rommers and colleagues (2013) and.90 in Piai and colleagues (2014; see table 1, p. 148). Thus the expected final words were indeed predictable. The expected and anomalous sentence-final-words of each sentence frame ("koekje" and "muisje" in this example) were matched in logarithmically transformed frequency of occurrence (M expected = 2.023, SE expected =.109; M anomalous = 2.008 ms, SE anomalous =.105; t (318) =.098, p =.921), number of phonemes (exact match), number of syllables (in all but 3 cases), gender, and number (in all but 2 cases).

The 164 sentences were presented intermixed for recording in a sound-shielded booth by a male native speaker of Dutch. The two versions of each sentence (expected and anomalous) were never closer than five sentences apart. We used whole sentence recordings rather than splicing the expected and anomalous sentence-final-word onto each sentence frame. The duration of the sentences did not differ between conditions (M expected = 4076 ms, SE expected = 102.23; M anomalous = 4111 ms, SE anomalous = 102.18; t (318) = -.2431 p =.808), nor did the duration of the sentence-final-word (M expected = 559 ms, SE expected = 9.8; M anomalous = 571 ms, SE anomalous = 12; t (318) = -.7657, p =.444).

Two experimental lists were created (i.e., sequences of trials) to counterbalance predictability. Each list included only one version of a given sentence frame (ending in either the expected or the anomalous word). For both lists a comparison was made within and across lists to make sure that sentence duration and sentence-final-word duration did not differ (see table in appendix I for a detailed description).

These two lists were further crossed to counterbalance task (plan, no-plan). Thus, a total of 4 lists were used in the exposure phase, such that each participant was presented with stimuli from only one list, and each sentence frame was presented once in each of the 4 lists in only one of the 4 possible combinations: expected no-plan, expected plan, anomalous no-plan, and anomalous plan. The sentence frames were presented in the same order to the participants of all lists. Only Predictability (expected vs. anomalous sentence-final-words) and Task (plan, no-plan) differed.

Eighty-two pictures were selected from the database of Severens, Van Lommel, Ratinckx, and Hartsuiker (2005) to be used for the naming task (for plan trials; see appendix L for a full listing of the pictures). Two of these were used on practice trials. The sentences were combined with a picture, whose name was semantically and phonologically unrelated to the sentence-final word. For no-plan trials, each sentence was presented together with a meaningless line drawing. Pictures were presented as black line drawings on a white background (see Figure 1), scaled to fit into frames of 300 by 300 pixels (4.5° of visual angle from the participant's position).

The materials for the recognition memory test consisted of the 160 sentence-final words of a given exposure phase list (see appendix K) and 80 lures (240 trials in total; see appendix M), which were drawn from the sentence-final words of the other exposure phase lists. As such, for recognition memory test lists 1 and 3 lures were drawn from the sentence-final words of exposure phase lists 2 and 4 and vice versa. That is, as described above not all participants heard the same sentence-final-words during the exposure phase and as a result the same was true for the recognition memory test lists. The 80 lures were randomly picked out of the 160 possible lures. These 80 lures were used for all participants that were administered a given list[1]. Additional testing made sure they did not differ significantly from the 160 targets in spoken duration (Lists 1&3: $M_{targets}$ = 749 ms, $SE_{targets}$ = 12.96; M_{lures} = 742 ms, SE_{lures} = 18.53; t (238) = -.282, p =.777; Lists 2&4: $M_{targets}$ = 767 ms, $SE_{targets}$ = 13.72; M_{lures} = 736 ms, SE_{lures} = 18.91; t (238) = -1.315, p =.189) and lemma frequency (Lists 1&3: $M_{targets}$ = 1.98, $SE_{targets}$ =.106; M_{lures} = 2.06, SE_{lures} =.145; t (238) =.422, p =.672; Lists 2&4: $M_{targets}$ = 2.04, $SE_{targets}$ = 0.108; M_{lures} = 2.03, SE_{lures} =.144; t (238) = -.025, p =.98). Two more lures were added to be used in the practice of the recognition memory test, which also included two of the practice items from the exposure phase. The order of the targets was kept constant between the exposure phase and the recognition memory test, while zero to 4 lures intervened between successive targets. The recording of the words (targets and lures) was done by the same speaker that had produced the sentence materials. No information was given to the speaker regarding the type of stimulus recorded (target or lure).

Procedure. The experimenter prepared the participants for the EEG recording. Subsequently participants were seated in a soundproof booth in front of a computer screen. They were given a sheet of paper with written instructions for the exposure phase task. Participants had to name all pictures correctly and as quickly as possible, while also listening to the sentences. They were also instructed to try to blink as little as possible and blink -if necessary- only when the three asterisks appeared on screen. If instructions were clear, the four-trial practice session started. After the practice part participants could ask for clarification. Once the task was clear to the participants the actual exposure phase started (160 trials, 8 blocks of 20 trials each). There were short breaks to rest between the blocks. The exposure phase took approximately 30 minutes to complete.

Figure 1depicts the trial structure in the exposure phase. Every trial started with a short beep tone followed by a fixation cross for 750 ms, after which the spoken sentence was heard. At sentence-final-word onset either a picture or the meaningless line-drawing appeared on screen and stayed there for 250 ms. The visual stimulus was then replaced by a fixation cross for 1750 ms. Finally a series of three asterisks (***) appeared on screen to signal that the participant could now freely blink (3000 ms). After the exposure phase, participants went on to complete an intervening task (the number-letter task, the details of which are not described here). This took approximately 20 minutes.

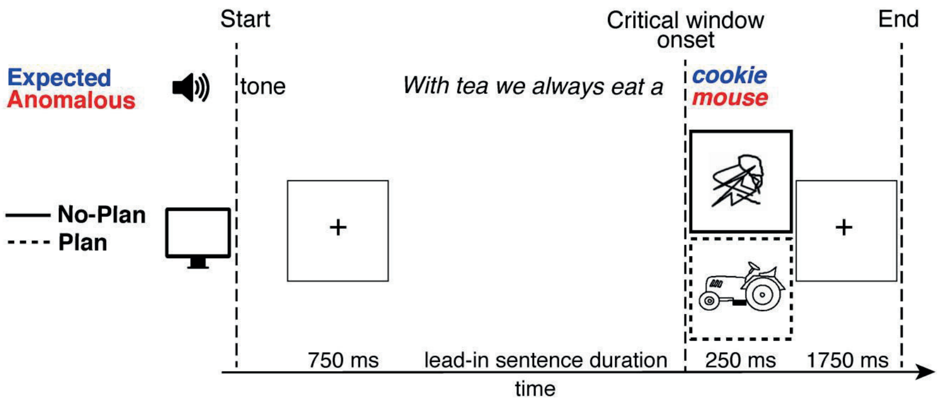

Figure 1. Trial structure in the exposure phase. The expected ("cookie") or anomalous ("mouse") sentence-final-word was heard in sync with the presentation of the meaningless line-drawing (no-plan task) or with the to be named picture (plan task).

For the recognition memory test participants read the instructions on a sheet of paper before the practice session started. Any questions were clarified before the recognition memory test session started. Then participants were presented with the sentence-final-words of the exposure phase intermixed with lures and were asked to indicated for each spoken word, whether they had heard that word in the exposure phase or not. This binary decision was documented with the help of a button box on which participants had to press the "new" or "old" button as quickly as possible. "New" and "old" response buttons were counterbalanced (left-right, right-left) across participants.

At the beginning of each trial participants saw a fixation cross for 1000 ms; then the single word was heard. Participants had three seconds to make a decision during which a fixation cross was displayed on screen. Then the fixation cross was replaced by three asterisks to allow for some blinking time (2000 ms). The 240 recognition memory test trials were presented in 12 blocks. At the end of each block the participant could rest. In total, it took approximately 40 minutes to finish this experiment part.

EEG Recording. The signal was amplified with BrainAmp amplifiers, filtered on-line using a 0.016 - 100 Hz band-pass filter and sampled at 500 Hz. Sixty active Ag/AgCl electrodes were mounted equi-distantly in an elastic cap (ActiCap) while 2 more electrodes were placed above and below the left corner of the mouth to monitor muscle activity from articulation onset. The horizontal EOG was measured by electrodes placed on the outer canthus of each eye (cap mounted electrodes). The vertical EOG was measured by electrodes on the supra-orbital ridge of each eye and a cap mounted electrode above each eye. The ground electrode was placed on the forehead. Electrode impedance was kept below 10 kΩ.

EEG Analysis. The preprocessing of the signal was done in Brain Vision Analyzer (version 2.0.2.5859). For five participants voltages on one electrode had to be interpolated from surrounding electrodes. First, all electrodes were re-referenced off-line to averaged mastoids.

Then, bipolar horizontal and vertical EOG were computed. Next, a similar difference score was computed for the mouth using the two electrodes next to the mouth corner. The continuous signal was filtered using a Butterworth Zero Phase Filter (Low Cutoff: 0.05 Hz, High Cutoff: 20 Hz) and segmented into epochs starting 200 ms before sentence-final-word onset (note, as described above, sentence-final word onset coincided with picture onset) and ending 700 ms after sentence-final-word onset. Following segmentation the signal was baseline corrected to a 150 ms window preceding sentence-final-word onset. Epochs with eye blinks were detected using a semiautomatic artifact rejection procedure. The same procedure was followed for horizontal eye movements and any other muscle artifact. After artifact rejection an average of 37.1 (range 31 to 40), 37.6 (range 30 to 40), 34.6 (range 29 to 40) and 35.4 trials (range 29 to 40) per participant were left in the no-plan correct, no-plan unrelated, plan correct and plan unrelated tasks respectively. Trials in which participants made a naming error were kept in the data, because for this study it was important that people were planning while listening, irrespective of whether what they produced was correct or not (and removing errors from the data would result in too few trials in the plan task).

For statistical analyses the output from Brain Vision Analyzer was read into Fieldtrip (Oostenveld, Fries, Maris, & Schoffelen, 2011). First, plan trials with a response time faster than 400 ms and no-plan trials to which participants might have tried to say something by mistake were excluded from the data. An average of 37.04 (range 31 to 40), 37.54 (range 30 to 40), 33 (range 26 to 40) and 33.6 trials (range 25 to 38) per participant were left in the no-plan correct, no-plan unrelated, plan correct and plan unrelated tasks respectively. To compute the Event Related Potentials (ERPs) the trials of each condition were averaged within participant for each condition.

Non-parametric (permutation based) statistics across the entire -200 to 700 ms epoch (without pre-selecting a time window) were used to test whether the factors Task and

Predictability modulated the size of the N400. First, the size of the significant clusters was established for the observed data (clusters spanning both the spatial and temporal domain). Next 1000 random permutations (trials randomly assigned to condition labels) were generated and the maximum cluster was calculated for each permutation. The distribution of permutation-based cluster sizes was used to determine significance of the actually observed data (see Maris & Oostenveld, 2007; Maris, 2012, for more detail).

Results

Exposure phase

Behavioral Data. Response speed was evaluated with linear mixed-effects regression models in R (version 2.14.2; The R foundation for statistical computing; lme4 package; Bates et al., 2015). A design-driven approach was adopted and the random structure of the models was determined following the suggestions of Barr, Levy, Scheepers, and Tily (2013). To evaluate interactions we used a backward elimination procedure that examined whether removing a specific interaction in the model improved model fit. Step by step, interactions that turned out to be non-significant were left out from the equation. All predictors justified by the design were kept in the models. In addition, a null model that included only the random structure of the models was established. Model comparison was made using log-likelihood ratio tests.

For evaluating response speed, naming errors were removed from the data set (19.5 %). In addition, responses faster than 300 ms or slower than 2000 ms were removed from the analyses (1.14 % and 0.2 % of the data respectively), as were data points with a value further than two standard deviations away from the participant mean (4.7%). Reaction Times (RTs) were measured from picture onset on (note that picture onset was synchronous to the onset of the sentence-final-word). Median response speed was 1026 ms and mean response speed 1051 ms. Predictability (expected vs anomalous, fitted as 1 and -1 respectively) and List (centered) were included as fixed factors in the model. The random structure of the model included random intercepts for participants and items (here, pictures) as well as a random slope (by participants) for predictability[2]. Listening to expected sentence-final-words led to faster picture naming (1032 ms) than listening to anomalous sentence-final-words (1071 ms; $b = -20.66$, $p <.001$).

EEG Data. Figure 2 features a representative subset of 11 electrodes depicting the grand-average ERPs for the expected and anomalous sentence-final-words in the two tasks (plan, no-plan). To investigate the interaction between the factors Predictability and Task, two matrices, each representing the spatiotemporal differences (expected vs. anomalous in the no plan task and expected vs. anomalous in the plan task) were computed. These difference matrices were compared to each other with the cluster-based permutation approach mentioned above (i.e., this approach has no a priori assumptions on the time-window involved). The analysis revealed a significant interaction between Predictability and Task (p =. 042). The difference was associated with a cluster that started 438 ms after sentence-final-word onset and lasted until 698 ms. The interaction was further evaluated by looking at the predictability effect in the no-plan task and in the plan task separately. In accordance with the statistical analysis for the interaction effect, no a priori assumptions were made for the time-window in which an effect might arise[3]

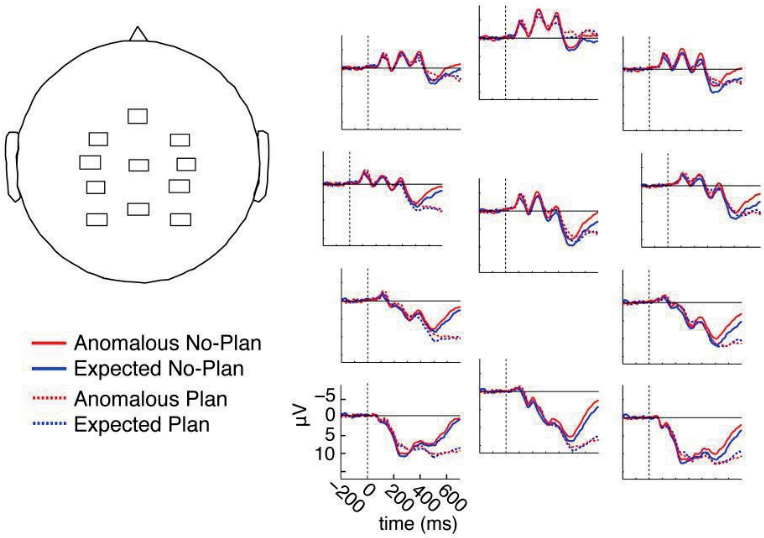

Figure 2. Event-Related potentials (ERPs) of a representative subset of 11 electrodes (locations are indicated on the head) timelocked to the sentence-final-word (also picture onset) as a function of Task (plan, no-plan) and Predictability (expected, anomalous) in the exposure phase. X-axis is time in milliseconds; Y-axis is amplitude in microvolt (µV).

The cluster-based permutation test for the main effect in the no-plan task revealed a significant difference between the expected and anomalous sentence-final-words (p =.01), with anomalous sentence-final-words inducing a larger negativity than expected sentence-final-words. No difference was observed in the plan task (p =. 478). The cluster in the no plan task was detected in the time-window between 368 ms and 698 ms after sentence-final-word onset. Figure 3 displays the scalp distributions of the interaction effect (left topographical map) and of the predictability effects in the no plan (middle topographical map) and in the plan task (right topographical map).

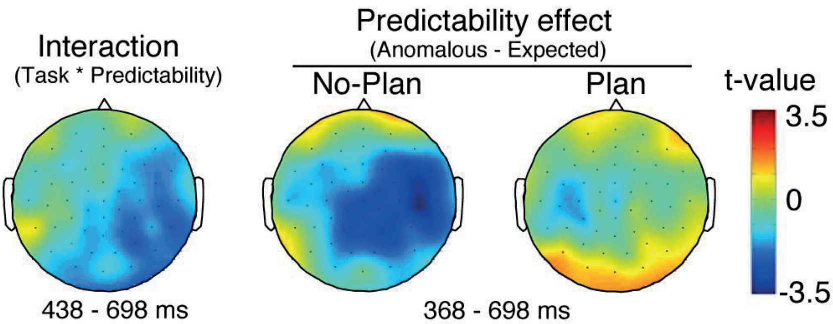

Figure 3. Scalp distributions of the interaction effect (left topographical map), of the predictability effect in the no-plan (middle topographical map) and in the plan task (right topographical map). Note that here the coloring represents t-value and not microvolts. The maps have been uniformly scaled. The time interval corresponds to the time interval in which the significant cluster was found.

A potential concern may be that an early planning-related component could have partly occluded a potential N400 effect in the plan task. It should be noted, of course, that the manipulation of Task and Predictability was fully orthogonal, which should render this effect minimal. Nonetheless, an additional analysis again with cluster-based permutation testing assessed the main effect of task. Plan and no-plan signals were compared and one positive cluster (p =.003, 436-698 ms) and one negative cluster (p =.019, 118 - 348 ms) were detected. The positive cluster was due to the plan task resulting in a larger positivity than the no-plan task, while the negative cluster resulted because the plan task induced a larger negativity compared to the no-plan task. Given its timing, the difference in the signal resulting in the late positive cluster probably reflects the difference between articulation preparation (plan task) and no articulation preparation (no-plan task). This finding is important, because it demonstrates that the absence of the N400 effect in the plan task was not because overlapping production planning components obscured the effect, since the planning effect seems to onset much later than the N400 effect (436 ms vs. 368 ms). The observed negative cluster starts very early in time and has a posterior (rather occipital) distribution (see figure 4). This might reflect covert orienting of attention (Hillyard & Munte, 1984) linked to visual target recognition, by which a novel picture would signal the plan task, while the standard shape would signal the no-plan task.

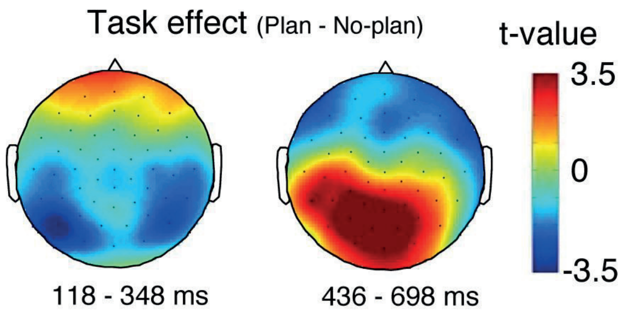

Figure 4. Scalp distributions of the planning effect. The left topographical map depicts the negative cluster and the right topographical map the positive cluster. Note that here the coloring represents t-value. The maps have been uniformly scaled. The time interval corresponds to the time interval in which the significant cluster was found (118-348 ms for the negative cluster and 436-698 for the positive cluster.).

In an additional more liberal analysis, we re-computed the permutation test for a restricted time-window 300 to 700 ms from sentence-final-word (and picture) onset. Once more, we could not locate a significant N400 amplitude difference between expected and anomalous sentence-final-words in the plan task ($p = .306$). Analysis in the exact same time-window (300 to 700 ms) for the no-plan task did reveal a significant negative cluster (p =.01), thereby further confirming the presence of an N400 effect in the no-plan task.

Recognition memory test

Behavioral Data. We assessed the recognition memory accuracy of the participants as a function of Task (plan, no-plan) and Predictability. Mean percentage correct was 63% and 60% for plan expected and plan anomalous, and 67.6 % and 67% for no-plan expected and no-plan anomalous respectively. Recognition memory accuracy was quantified by computing d', which is better suited to pick up actual differences between conditions. This is the sensitivity signal detection index (Macmillan & Creelman, 2005). Figure 5 depicts mean d' values by task and predictability. A two-way (plan, no-plan; expected-anomalous) repeated measures ANOVA showed a significant effect of task (F_1 (1, 23) = 92.466, *p <.001*, η^2 =.801) and of predictability (F_1 (1, 23) = 6.003, *p =.022*, η^2 =.207). Items of the no-plan task were remembered better ($M_{No\ plan}$ =.913, $SE_{No\ plan}$ =.095) than items of the plan task (M_{plan} =.431, SE_{plan} =.077). Moreover, expected sentence-final-words were remembered better ($M_{Expected}$ =.745, $SE_{Expected}$ =.072) than anomalous ones ($M_{Anomalous}$ =.599, $SE_{Anomalous}$ =.102). The interaction of task with predictability was also significant (F_1 (1, 23) = 4.987, *p =.036*, η^2 =.178). Paired-samples t tests on the participant means (t_1) indicated that the predictability effect was significant only in the plan (t_1 (23) = 4.056, *p <.001*), and not in the no-plan task, t_1 (23) =.507, *p* =.617. Thus participants were better at remembering the expected sentence-final-words than the anomalous ones, but only for those items that had been presented while they were planning a picture name. (*M* =.2489, *SE* =.061). For items that had been presented while participants had not been planning, there was no effect of predictability (*M* =.0442, *SE* =.087).

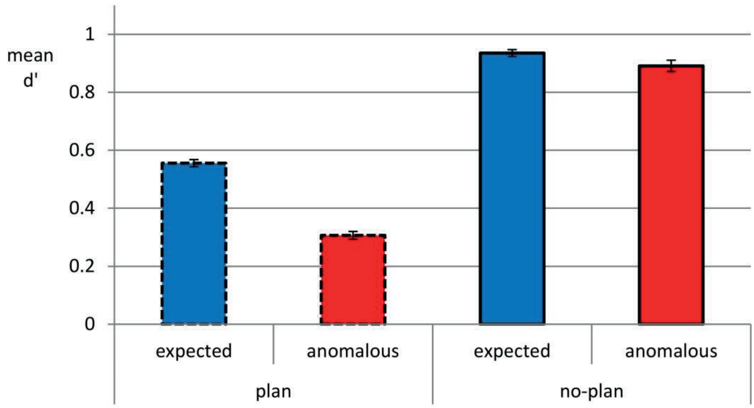

Figure 5. Mean d' value by task (plan, no-plan) and predictability in the recognition memory test.

For the analysis of decision speed outliers were set to two standard deviations from the participant mean (4.37 % of the data excluded). The auditory word offset served as the timepoint from which on decision latencies were measured. The initial model included Task (plan, no-plan; fitted as 1 and -1, respectively), Predictability (expected, anomalous; fitted as 1 and -1, respectively), response (new, old; fitted as -1 and 1, respectively), List (centered) and the interaction of task with predictability and response as fixed factors, a random intercept by participants and spoken word, and a random slope of task and predictability by participants and by spoken words (and the corresponding interactions). Note that here we evaluate the responses to actual targets and not to lures. The best fitting model included significant main effects of response ($b = -75.179$, $p < .001$, intercept = 823) indicating that participants were faster to make correct “old” decisions. Task ($b = 12.540$, $p = .066$), predictability ($b = 12.386$, $p = .150$) and list ($b = 5.924$, $p = .887$) did not contribute significantly to the model. A significant two-way interaction was observed between Task and Response ($b = 19.256$, $p = .003$). The interaction of task with response was further explored by running two additional models; one for the level "old" and for the level "new" of the factor "response". Task, predictability, and list as well as the interaction of task with predictability were entered as fixed factors in the initial model. The random structure was identical to the model used for evaluating the decision speed overall. The analysis showed a significant contribution of task (plan, no-plan) only when labeling an item as "old" ($b = 30.185$, $p < .001$, intercept = 754.66). In particular no-plan items were labeled "old" significantly faster than plan items ($\text{No-Plan-Old}_{\text{mean}}$ = 711 ms, $\text{No-Plan-Old}_{\text{standard deviation}}$ = 420; $\text{Plan-Old}_{\text{mean}}$ = 758 ms, $\text{Plan-Old}_{\text{standard deviation}}$ = 427). This was not the case when labeling an item "new" ($b = -11.858$, $p = .242$, intercept = 915.94; $\text{No-Plan-New}_{\text{mean}}$ = 918 ms, $\text{No-Plan-New}_{\text{standard deviation}}$ = 490; $\text{Plan-New}_{\text{mean}}$ = 907 ms, $\text{Plan-New}_{\text{standard deviation}}$ = 487).

Discussion

The present study investigated to what extent the quality of listening to speech deteriorates when people are engaged in speech planning and whether such performance loss is mediated by the predictability of the sentence-final-word. In the exposure phase participants heard sentences that ended in a semantically expected or semantically anomalous word. On half of those trials the sentence-final words were presented while participants were also probed to name a picture (i.e., they were engaged in speech planning during the presentation of the sentence final word), while on the other half of the trials they did not have to name a picture. Reaction times for picture naming and ERPs reflecting cortical processing during the processing of the sentence final words were analysed. In a subsequent recognition memory test participants were presented with the sentence-final-words from the exposure phase along with a set of lures, and were asked to indicate which of these words they had heard before. For the recognition memory test memory performance and decision times were analysed.

Semantic processing is affected by concurrent speech planning.

The ERP results revealed that when participants were planning to name a picture, the semantic manipulation did not evoke a typical N400 effect, which was observed when participants did not prepare a picture name. The absence of an N400 effect in the plan task is consistent with the idea that planning interfered with thorough semantic processing of the speech that participants heard. An influence of planning speech on thorough semantic processing has also been reported in a very recent study by Bögels and colleagues (2018). This study is described in more detail below. In addition, however, results of the subsequent recognition memory test also revealed influences on behavior. That is, participants had poorer memory for the sentence-final-words that they had heard while they were planning to name a picture (as compared to those that they heard without concurrent planning). This finding is compatible with the idea that recognition memory performance drops when attention is divided during encoding (e.g. Fernandes and Moscovitch, 2000). The decision times in the recognition memory test further supported this notion; as the interaction of task (plan, no-plan) with response (old, new) revealed that correct "old" decisions were made faster for no-plan than for plan items. Thus, both neural online measures and behavioral offline measures seem to suggest that planning while listening to the sentence-final-word disrupted semantic processing of that word [4].

Although the possibility cannot be fully excluded here, we deem it unlikely that the N400 effect was not detected in the plan task because of overlapping production-planning components in the EEG signal. The additional analyses comparing the EEG signal of the plan task to that of the no-plan task indicated that the strongest effect of production planning emerged later in time than the N400 effect in the plan task. Moreover, in the study by Bögels and colleagues (2018), which similarly to our findings reports a positivity starting around 500 ms for the early-planning condition and somewhat earlier for the late-planning condition, and which relates that to response planning, an N400 could be registered for both early- and late-planning. Thus, the positivity ascribed to the response planning processes did not obscure the N400 effect in any of the conditions (we further elaborate on this issue below). Note that both studies used cluster-based permutation testing to test for statistical significance. From that we can conclude that our analyses method is probably robust enough to have picked up an N400 in the plan task, if there was one. Yet, even limiting the time-window of analysis to 300-700 ms after sentence-final-word onset did not reveal an N400 effect for the plan task. Consider also that in our study predictability and speech planning were manipulated orthogonally. Thus both expected and anomalous items were equally affected by planning.

In addition, it is unlikely that the N400 effect was not detected in the plan task because of articulation artifacts in the EEG signal, since only the time-point up to 700 ms after sentence-final-word (and picture) onset was evaluated, while mean response speed was 1051 ms and median response speed was 1026 ms. Although these naming latencies may seem somewhat long (Indefrey, 2011), in the current experiment participants first had to decide whether they had to name or do nothing in the given trial, and only on a naming trial actually initiate articulation. However, the longer latencies decreased the likelihood of overlap between the N400 and articulation based artifacts.

But why is it that the current study did not find an N400 effect for sentence-final-words heard while planning, whereas Bögels and colleagues (2018) report an effect for both the early- and the late-planning condition? Alike to the study reported in this chapter, their study investigated the impact of planning on comprehension processes, as reflected by the N400 component. But rather than contrasting planning with no-planning, they compared early- to late-planning. For a start it is important to note that although Bögels and colleagues (2018) did not observe a reduced N400 effect in their experimentally induced early vs. late planning manipulation, the authors did observe a reduced N400 response when restricting the analyses to participants that had responded fast (i.e., those with short inter turn intervals, as compared to so-called slow responders).

Moreover, in the sentences used in the present study there was complete time-overlap of the point at which planning could initiate (plan task) and of the point at which the anticipated word could be perceived, namely the sentence-final-word. In the study by Bögels and colleagues (2018) the point at which a word could be anticipated was either before planning could initiate (that is, before hearing the critical word that would allow the participant to choose the correct object-name out of two picture-candidates), or after planning could initiate (that is, after hearing the critical word that would allow the participant to choose the correct object-name out of two picture-candidates). The participant would see two pictures on screen (banana and apple) and after a short delay would hear one of the following example sentences: a) "Which object is **curved** and is considered to be a type of fruit?" (**early-planning**, expected), b) "Which object is **curved** and is considered to be healthy?" (**early-planning**, unexpected), c) "Which object is considered a type of fruit and is **curved**?" (**late-planning**, expected), d) "Which object is considered healthy and is **curved**?" **(late-planning**, unexpected). Thus the participant would have to produce the response "banana". Crucially, the point at which the anticipated word "fruit" or the unanticipated word "healthy" could be perceived, never coincided with the point at which planning could start. And the keyword here is "start".

This absence of complete time overlap between the point of perceiving the anticipated or unanticipated word, and perceiving the critical word that would trigger the plan response, might have allowed for the N400 effect to survive the early-planning condition. Maybe hearing "healthy" (unexpected condition) as opposed to "fruit" (expected condition) at time point 2 can easily be processed (resulting in an N400), even when already planning the word "banana" upon having heard "curved" at time point 1, because from time point 1 to time point

2 there was sufficient time to prepare the most effortful parts of production planning for the word "banana". Given that following the design of the trial structure actually participants might already have been buffering the two candidate responses ("banana" and "apple") planning might be even easier to do.

Moreover, in the study by Bögels and colleagues (2018) the anticipation seems to be of a somewhat different kind compared to the current study: the participant might anticipate the word "fruit", because at the beginning of the trial she was shown two pictures of, for example, a banana and an apple. In addition though, one could argue that upon hearing the word "curved" (which is meant to be the planning initiation trigger) she also experiences a kind of associative retrieval of the word "banana", which is also the one she has to utter as a reply to the confederates question. Thus, at least the kind of anticipation seems to differ between the two studies. In the present study we tested the anticipation resulting from the preceding sentence context that should in most cases lead to anticipating one word, while in the study by Bögels and colleagues (2018) the anticipation was mainly driven by the pictures presented to the participant at the beginning of each trial (e.g. seeing the pictures of a banana and an apple led to anticipation of "fruit") and partly by this associative retrieval of the answer, triggered by the critical word (e.g. "curved" triggers the response "banana").

Yet again, another factor might be at work that stems from the instructions. In the present study the instructions given to the participants were to name the pictures as quickly as possible but also listen to the sentences, because later they would have to do something with these. In this sense, listening was important but not as vital as in the study by Bögels and colleagues (2018), where one had to have heard the critical word to pick the correct object to name. This being a more interactive setting than the present study, might have prompted participants to switch their attention away from listening at a later point in time than participants in the present study. As a result an N400 was registered for both early- and late-planning in the Bögels and colleagues (2018) study, while in the current study no N400 was registered for the plan task, because attention was shifted away from listening and towards planning much faster in the current study. One argument against this interpretation might be that despite the absence of an N400 effect in the plan task of the present study, recognition memory performance was still mediated by predictability, which implies that the sentence-final-words were none the less processed to some degree. We return to this point in the following section. Yet, based on the current data we cannot completely exclude this possibility and definitely addressing this issue would require additional testing. All in all, looking into factors affecting the time point of switching attention between listening and planning seems to be a promising future research avenue.

Predictability partly mediates the cost of concurrent speech planning.

A second pattern that was observed in the recognition memory test was that for the items that were presented during planning in the exposure phase, participants showed a gain in memory performance from predictability, with expected sentence-final-words recognized more often than anomalous ones. This finding demonstrates that part of the performance loss in the recognition memory test due to planning, could be mediated by predictability. No such gain in memory performance from predictability was registered in the no-plan situation.

Predictability did not result in a significant increase in memory performance when participants did not have to plan while listening, mainly because in this situation (no-plan task) the anomalous words were recognized more often than in the plan task. Maybe participants had more time or more resources to spend on processing the anomalous sentence-final-words when not planning than when planning. The robust N400 effect in the no-plan task suggests that indeed anomalous items were more thoroughly processed in the no-plan task. As a result they could probably better encode the word heard. Thus, in the no-plan setting anomalous sentence-final-words were remembered equally well with expected ones and predictability could not add to memory performance.

Contrary to that, in the plan task, the sentence-final-words could not be processed thoroughly, presumably due to resource competition in this perception/production setting. As a result overall memory performance dropped and we could not detect an N400 effect. However, because planning only began around the end of the sentence, relatively intact memory of the sentence context may have led participants to correctly respond “old” to the expected sentence-final word. In this case the sentence-final-word, even though not thoroughly processed due to planning, could be associated to a perceived sentence context leading to a correct "old" response, when encountered in the recognition memory test [5]. Such a pattern providing a strong basis for an "old" response could not emerge for the anomalous sentence-final-words, which -like the expected sentence-final-words- were also not well processed, but at the same time could also not easily be linked to a perceived sentence context.

Our study is not the only one reporting the absence of an N400 effect, while still tracing evidence of semantic processing via behavioral measures. The plan no-plan task manipulation in our study is comparable to the attended-unattended manipulation in the study by Bentin and colleagues (1995). They found no N400 for the unattended items and we did not find an N400 for the plan task, i.e. for the situation in which attention was probably (at least partly) driven away from listening due to planning. Moreover, attended items were remembered better than unattended in a recognition memory test, like in our case no-plan items were remembered better than plan items. Yet, both attended and unattended items produced comparable repetition memory effects in a lexical decision task. They conclude that such a pattern demonstrates that despite the absence of an N400 in the unattended condition, both item types (attended and unattended) were still semantically processed and activated semantic representations. That is, at least some aspects of semantic processing took place even for the unattended words, but these were not picked up by the N400. In a parallel to that we argue that even though semantic processing was severely disrupted by planning, still some aspects of semantic processing took place thereby resulting in the memory benefit for expected compared to anomalous sentence-final-words in the plan task. Traces of semantic processing in the absence of an N400 effect have also been reported by Kellenbach and Michie (1996), who report that even though unattended words were recognized less often than attended words, unattended words were still better recognized than chance level.

How can one further explain why we did not pick up a difference between expected and anomalous items in the no-plan task, while the existing literature reports a difference in memory performance between semantically related and unrelated items (Besson et al., 1992; Corley et al., 2007; Kellenbach & Michie, 1996; MacGregor et al., 2010; Olichney et al., 2000; Phillips & Lesperance, 2003; Wlotko et al., 2012)? One additional explanation may have to do with the type of material used in our study compared to the cited studies. For the semantically anomalous condition we picked sentence-final-words that were not in any way plausible as sentence endings of these sentence-frames. While for example Corley and colleagues (2007) and MacGregor and colleagues (2010) used unpredictable but still plausible sentence-final-words, which may have a greater probability of being integrated into a sentence or message level representation in memory. The material we used in chapter 4 introduced a similar rather subtle kind of predictability manipulation, where sentence-frames were either constraining or non-constraining but still plausible. There we observed no modulation of memory performance by predictability while planning. It seems thus that the kind of predictability manipulation introduced - and this again links to Bögels and colleagues (2018) - has an impact on whether it will interact with planning or not.

It was not only memory performance that gained from predictability in the plan task. As our response speed data in the exposure phase suggest, pictures were named faster when paired to an expected sentence-final word. That means that predictability of the sentence-final-word also facilitated timely planning of an unrelated to it picture-name. This likely stemmed from later (or different) processes not captured by the N400. Bögels and colleagues (2018) also report that in the condition where planning overlapped with listening (early-planning condition) participants were faster to respond, when having heard an expected compared to an unexpected word.

Implications for the conversational setting and possible sources of interference.

Our findings seem to imply that the short pauses usually reported between interlocutors' turns (e.g. Beattie & Barnard, 1979; Bosch, Oostdijk, & Boves, 2005; Stivers et al., 2009) might at times come with a cost for listening. We demonstrated that semantic processing can be affected by such a dual setting of planning speech while still listening to speech. The cost of dual tasking was evident both online (EEG recordings) and offline (Recognition Memory Test).

Given that a number of studies reviewed in the introduction of this article reported that N400 amplitude remained unaffected by dual tasking and attentional manipulations of several kinds, it is rather surprising that a task involving single-word planning while listening should diminish the N400 effect to that degree. One would think that overlapping production and perception processes are that common, given the short pauses between turns, that our participants should have enough practice with that overlap not to be affected. Previous research demonstrating that the neuronal structures (regarding semantic, lexical and syntactic processing) involved in comprehension and production are at most common (Menenti et al., 2011) and that a switching cost might be the underlying factor impacting semantic processing in dual task situations with high temporal overlap (Vachon & Jolicoeur, 2011; Vachon & Jolicœur, 2012), could possibly explain part of the cost registered in our study, but the current study does not offer the tools necessary to further elaborate on this.

The idea of switching costs as the source of interference seems to be compatible with the findings by Bögels and colleagues (2018) according to which fast responders - compared to slow responders- allocated their attentional resources faster towards planning and away from comprehending (possibly away from anticipating). This finding in turn is compatible with the finding by Hohlfeld and Sommer (2005) that high temporal overlap between two tasks and increased task load reduce the N400 amplitude. The authors linked this amplitude reduction to the need to share resources between tasks (Sommer & Hohlfeld, 2008; see also Luck, 1998). A similar amplitude decrease reported in Hohlfeld and Sommer (2015) was interpreted in the framework of the attentional sensitization model by Kiefer and Martens (2010). This model posits that attention can be driven by internal goals or task representations thereby enhancing perceptual sensitivity for target stimuli. Thus depending on how much the focus is on the listening or the planning side, the outcomes would be variable.

Irrespective of the possible mechanisms behind it, this kind of cost demonstrated by our study might be a cost an interlocutor could tolerate for the sake of a short inter-turn interval. Such an idea could match the good enough approach to language processing (Ferreira, Bailey, & Ferraro, 2002; Ferreira & Patson, 2007). Most likely there are a number of other linguistic dimensions for which a cost could be detected. These could span from pragmatic aspects being missed out on, up to syntactic errors, and can even reach to issues concerning the overall episodic memory of a conversational interaction.

Conclusion

To sum up, we set out to investigate whether single word planning while listening to speech would affect listening quality not only offline (behavioral measures) but also online (EEG recordings) and whether any cost seen could be alleviated by predictability. In accordance with our expectations, participants' memory was best when they had not been planning while listening to the sentence-final-word. The weakened memory performance was improved when the sentence-final-word that was heard while planning was expected. The absence of an N400 effect in the plan situation also suggests that the sentence-final-word was processed less thoroughly when planning. Even though our study material does not qualify as a conversational setting, one can conclude that, at least at the level of processes in such a simplified situation, interlocutors likely have to effectively allocate their resources between planning and listening to produce a timely turn.

Acknowledgments

The authors thank Vitória Piai for valuable assistance in statistically analyzing the EEG data and Hanna de Groot for testing some of the participants.

Footnotes

[1] Note that given that the same 80 lures were used for all recognition memory test lists, some of the lures heard could actually be the items that could have been predicted by the participants during listening to the sentence frames, but were never heard, because the anomalous sentence-final-word was heard instead. To even out any effect of that we made sure that this was the case in a comparable number of cases across the recognition memory test lists: Out of these 80 randomly picked lures, 44 (list 1), 36 (list 2), 44 (list 3) and 36 (list 4) lures could have been predicted in that sense during the exposure phase.

[2] The random structure is described once for each dependent measure. Only if convergence problems forced us to alter it in any way it is mentioned again. When not explicitly stated, the random structure described at the beginning of each section applies to all the models of this dependent measure.

[3] The time-windows describing the detected clusters should not be interpreted as the exact on- and offsets in time for the effect under investigation, but rather as indicators of the most likely contributors (Maris, 2012).

[4] Based on the design of our study we cannot infer exactly how semantic processing might have been affected, even though one could assume that the observed effects are linked to difficulties in incorporating the heard item into an episodic memory trace, while planning.

[5] In a questionnaire administered to the participants after the completion of the study, seven out of twenty-four reported that occasionally they could decide that an item was "old", because they could remember the specific sentence context.

Chapter 6

Summary and conclusions

Summary of the results

Research on the timing of turn-taking - the act of switching between the roles of being a listener and being a speaker in conversation- suggests that interlocutors are very fast at taking over turns (e.g. Bosch, Oostdijk, and Boves, 2005; Stivers et al., 2009). Turn-taking timing has thus often been characterized as smooth and easy (e.g. Garrod & Pickering, 2004; Pickering & Garrod, 2004; Scott et al., 2009; Stivers et al., 2009). Current turn-taking models (Levinson & Torreira, 2015) assume that interlocutors would have to start planning their reply, while still listening, in order to adhere to this fast pace of turn-taking. This thesis investigated whether interlocutors indeed do plan ahead while still listening. The main focus was on whether the need to initiate speech planning, while still listening, impacts on listening quality.

Chapter 2 investigated whether the timing of speech and planning efficiency may affect turn-taking timing. In Experiment 1 of chapter 2 participants heard sentences, to which they had to respond by describing one of the two displays (left or right) that appeared on screen. The objects on the displays were configured in such a way that they would elicit a simple noun phrase followed by a complex noun phrase. For example, *"Het varken gaat naar de hamburger en de theepot"* (English translation: "The pig goes to the hamburger and the teapot"). To manipulate planning 'efficiency' the pictures on the displays were presented either in upright or in upside-down orientation. To manipulate speech planning onset one version of the auditory sentence had the critical information (left or right) early in the sentence (early timing condition: '*En het plaatje rechts/links, kan je beschrijven wat daar gebeurt?*' English translation: "And on the picture right/left, can you describe what happens there?"), while the other version had it at the end of the sentence (late timing condition: '*En kan je beschrijven wat er gebeurt op het plaatje rechts/links?*' English translation: "And can you describe what happens on the picture on the right/left?"). The results of this experiment revealed an impact of speech planning on turn-taking timing; participants were faster to respond when the critical information (left/right) was presented early, rather than late in the sentence. However, no impact of planning efficiency (picture orientation) on turn-taking timing was observed in Experiment 1.

Experiment 2 was identical to Experiment 1, only now the objects on the display were configured in such a way that they would first elicit a complex noun phrase followed by a simple noun phrase (e.g.*"Het varken en de hamburger gaan naar de theepot"* (English translation: "The pig and the hamburger go to the teapot"). As in Experiment 1, participants were faster to respond when the critical information (left/right) was presented early, rather than late in the sentence. In addition, however, and contrary to Experiment 1, an impact of planning efficiency was registered, such that in the late timing condition (i.e. when the critical information was presented only at the end of the sentence) having to plan the names of upside-down pictures resulted in longer response times compared to having to plan the names of upright pictures. This suggests that in Experiment 1 participants managed to quickly prepare the first simple noun phrase (regardless of orientation), and probably extended planning the subsequent complex noun phrase until after the onset of their response. In experiment 2 such an approach was impossible since the response required the complex noun phrase in initial position, which allowed the 'efficiency' effect to affect response times. The combined findings of chapter 2 thus suggest that the time at which planning starts directly contributes to turn-taking timing, while planning efficiency interacts with onset of planning.

The results of chapter 2, in combination with the literature on turn-taking timing (e.g. Beattie & Barnard, 1979; de Ruiter, Mitterer, & Enfield, 2006; Sacks, Schegloff, & Jefferson, 1974; Scott, Mcgettigan, & Eisner, 2009; Stivers et al., 2009; Wilson & Wilson, 2005) as well as recent studies (Bögels et al., 2018, 2015; Boiteau et al., 2014; Sjerps & Meyer, 2015), thus suggest that at the very least, interlocutors start planning their turn shortly before the turn-end of their interlocutor, and do not wait with planning their response until their interlocutor's turn is over. Hence, these findings further support the model put forth by Levinson and Torreira (2015). However, this observation also raises another question: If parallel processing of planning and listening do indeed occur, does that come at a cost for one or both of these processes? In other words: can we still listen carefully when we also initiate the planning of our response? And vice versa?

Chapter 3 investigated whether concurrent planning impacts negatively on the listening quality of heard words. Listening quality was evaluated offline by assessing recognition memory performance. The experiments in chapter 3 consisted of an exposure phase in which participants either engaged in word planning (planning task) or did not do so (no-planning task) while listening to single words. A recognition memory test followed the exposure phase. In Experiment 1 participants were not warned about the upcoming recognition memory test (incidental memory encoding mode), while in Experiment 2 they were (intentional memory encoding mode). In both experiments pupil dilation measures indicated that there was increased cognitive effort required for speech planning. For example, it was observed that pupil dilation was larger when planning speech (planning task) than when speech planning was absent (no-planning task). Furthermore, planning low frequency picture-names resulted in larger pupil dilation than planning high frequency picture-names.

Importantly, the cognitive effort put into concurrent speech planning was also found to affect spoken word processing: in both experiments, during the recognition memory test, participants were better at recognizing words heard in the no-planning than in the planning task. In addition, in Experiment 1 participants were also faster to (correctly) recognize words heard in the no-planning compared to the planning task. And when correctly recognizing items as “old”, pupil dilation was larger than when falsely recognizing items as “new” (experiments 1 and 2). Correct recognition of a word in the recognition memory test (albeit weakly) linked to larger pupil dilation during encoding in the exposure phase (experiments 1 and 2). These findings demonstrate that planning negatively affected listeners ability for subsequent recall. Interestingly, the impact of concurrent speech planning on spoken word processing could not be overcome, even when participants were intentionally trying to encode the heard words (Experiment 2). Yet, some trade-off was observed: shifting attention towards listening (intentional encoding in Experiment 2) significantly improved recognition memory performance in Experiment 2 (compared to Experiment 1), but at the same time negatively impacted on production performance. Participants made more production errors in Experiment 2, compared to Experiment 1. These findings clearly demonstrate a negative influence of production planning on the processing of heard words (recognition memory performance as an offline measure of processing quality). Moreover, they also demonstrate that listeners have the ability to strategically increase performance on one task at the cost of the other.

Given the short gaps observed in everyday turn taking, people might regularly be exposed to the kind of interference between listening and production planning that was demonstrated in chapter 3. However, everyday conversations rarely consist of single words. This raises the question whether listening to sentences may change this interplay. The context provided by a sentence may partly help to overcome any influences of production planning on listening quality. Like chapter 3, Experiment 1 of chapter 4 addressed the question on how well listeners cope with the dual demands of planning while listening. However, now participants listened to whole sentences rather than single words. Furthermore, an additional factor was introduced to test whether predictability of the sentence-final-word might improve memory performance in the recognition memory test. To this end, the sentence-final-words were embedded either in a constraining or in a non-constraining sentence frame. Participants engaged in word planning (planning task) or did not do so (no-planning task) while listening to sentences (exposure phase). A recognition memory test followed.

The results of Experiment 1 replicated the findings of chapter 3, but now in a sentence setting. Pupil dilation was larger during planning compared to not planning and larger when planning low- as compared to high-frequency picture-names (i.e., planning is effortfull). Moreover, sentence-final-words that had been heard while preparing to speak were recognized less often and slower (recognition memory test) than sentence-final-words heard without concurrent speech planning. However, this effect was not mediated by predictability; predictable sentence-final-words were not recognized more often than unpredictable ones, when planning while listening. Predictability did, however, affect naming speed in the exposure phase; naming a low frequency picture was faster when having just heard a predictable sentence-final-word compared to when having just heard an unpredictable sentence-final-word.

CHAPTER 6: SUMMARY AND CONCLUSIONS

The processing cost observed for listening (recognition memory performance and pupil dilation) in Experiment 1 as well as in Chapter 3, could have been caused by a difference in stimulus variability between the planning and no-planning conditions and not depend on the planning process itself. As such the observed costs for listening could arise from lower level processes, like picture recognition, rather than speech planning. To investigate this possibility a control experiment was conducted. To this end, Experiment 2 was identical to Experiment 1, only now participants were instructed to listen to the sentences and only passively view the pictures (i.e., pictures were *not* named in Experiment 2). No difference was observed in pupil dilation (exposure phase) and memory performance (recognition memory test) between the planning-control task and the no-planning-control task. The only difference that endured was that of faster decision speed when correctly identifying an item as "old", which was also accompanied by larger pupil dilation, compared to falsely identifying an item as "new" (recognition memory test). In comparing experiments 1 and 2 (exposure phase) it was found that task (planning vs. no-planning in Experiment 1 and control-planning vs. control-no-planning in Experiment 2) modulated pupil size and recognition memory performance in Experiment 1 but not in Experiment 2. This finding indicates that the observed effects in chapters 3 and 4 (Experiment 1) are indeed linked to planning processes and not to lower-level picture recognition processes. Chapters 3 and 4 thus demonstrate that concurrent planning negatively impacts the processing of heard words.

While these findings demonstrate that the parallel process of listening and planning comes at a cost, the conclusions based on chapters 3 and 4 were based on an offline measure of processing quality, namely recognition memory performance. Therefore chapter 5 focused on providing evidence that the impact of concurrent planning on listening quality can also be traced in real time. Rather than monitoring listening performance mainly offline, in chapter 5 I monitored listening performance online via electroencephalography (EEG), with an emphasis on the N400 component. As in chapters 3 and 4, participants either engaged in word planning (planning task) or did not do so (no-planning task). In chapter 5, however, participants listened to sentences that ended either in an expected or in a semantically anomalous sentence-final-word (exposure phase). Event-related-potentials (ERPs) elicited during listening to the sentence-final-words (exposure phase) revealed a substantial decrease in the amplitude of the N400 semantic anomaly-effect for items heard during speech planning compared to items heard when participants were not engaged in planning. This finding demonstrates that during speech planning, the encoding of heard speech is impaired, since it does not properly activate semantic representations. In accordance with chapters 3 and 4, recognition memory was poorer for sentence-final-words heard during the planning than during the no-planning task. In accordance with chapters 3 (Experiment 1) and 4, no-planning items were labeled "old" significantly faster than planning items. Contrary to chapter 4, the congruency manipulation (expected vs. semantically anomalous sentence-final-words) affected the quality of listening: Among the subset of sentence-final-words that had been presented in picture naming trials, those words that were expected given their preceding context were better remembered than the semantically anomalous ones. Congruency also affected response speed in the exposure phase, as listening to expected sentence-final-words led to faster picture naming than listening to anomalous sentence-final-words.

Conclusion

In conversation, interlocutors engage in turn-taking. Previous research has revealed that interlocutors are very fast in taking over a turn. The median gap between turns has been reported to be 330 ms (Ten Bosch, Oostdijk, and Boves, 2005; see also Beattie & Barnard, 1979; de Ruiter, Mitterer, & Enfield, 2006; Sacks, Schegloff, & Jefferson, 1974; Scott, Mcgettigan, & Eisner, 2009; Stivers et al., 2009; Wilson & Wilson, 2005). At the same time the speech production system is rather slow (Levinson, 2016); even single picture-naming can take 600 ms (Indefrey, 2011; Indefrey & Levelt, 2004), while preparing the first phrase of a sentence can easily take up more than a second (Konopka, 2011).

Given these speech planning preparation latencies, how do interlocutors manage to produce such a timely response? This question has been labeled the core psycholinguistic puzzle (Levinson & Torreira, 2015). In order to solve this puzzle interlocutors seem to have to engage in a dual tasking situation, wherein they have to start planning their reply while still listening to the incoming turn (Levinson & Torreira, 2015). This thesis aimed at investigating whether indeed interlocutors initiate speech planning while still listening. Moreover, this thesis investigated whether manipulating planning efficiency would interact with turn-taking timing.

Evidence from this thesis along with recent evidence from the literature (e.g. Bögels et al., 2015) suggests that indeed interlocutors initiate speech planning while still listening. That is, interlocutors engage in a kind of dual tasking situation, whereby they are planning while still listening. Yet, dual tasking typically has been shown to result in processing costs for one or both of the tasks (e.g. Baddeley, 1976; Becic et al., 2010; Kemper, Herman, & Lian, 2003; Kemper, Schmalzried, Herman, & Mohankumar, 2011; Lavie, Hirst, Fockert, & Viding, 2004; Lavie, 2005; Meyer &

Kieras, 1997; Pashler, 1984, 1994). Typically, increases in performance on one task are correlated with decreases in the secondary task (Somberg & Salthouse, 1982). Thus, in planning while listening interlocutors have to efficiently distribute their processing capacity between these two tasks.

How then do interlocutors perform in this dual tasking situation? Does listening performance drop, due to the need to prepare the speech? This thesis aimed at investigating how listening quality is affected by the need to prepare speech at the same time and whether there can be trade-offs between listening and preparing speech, depending on which task is given priority.

The findings of chapters 3, 4 and 5 indicated that indeed listening performance drops, when engaging in planning during listening. That is, there is a processing cost that results from this dual task situation. This finding brought about the question whether there are mechanisms which could help to lower the interference between listening and speech planning. One such candidate might be predictability, which has been demonstrated to affect processing speed, with predictable words being processesed faster than less predictable words (see for example Altmann & Kamide, 1999; Kliegl, Nuthmann, & Engbert, 2006; Traxler & Foss, 2000). Moreover, a number of studies have ascribed prediction a supportive role in either extracting the speech act of the incoming utterance as soon as possible, in order to initiate timely response planning (Levinson & Torreira, 2015), in facilitating comprehension of new information (Ferreira & Lowder, 2016), in anticipating a turn's end for a timely response initiation (Magyari et al., 2014, 2017; Magyari & de Ruiter, 2008), or in directly supporting a timely planning initiation via content prediction (Corps et al., 2018). As Levinson and Torreira argue in their turn-taking model, an essential part in solving the "core psycholinguistic puzzle" is to have predictive comprehension that will allow for extracting the speech act of the incoming utterance as soon as possible in order to inititate response planning. Rather than focusing on these aspects of prediction, this thesis focused on how sentence-final-word predictability might alleviate part of the processing cost related to preparing speech while still listening, thereby reducing the interference between the two.

In the following sections each of these questions is discussed, along with implications of the findings.

Do interlocutors plan their speech while listening?

The turn-taking model put forward by Levinson & Torreira (2015) assumes that interlocutors start to plan their response while still listening to the incoming turn. Chapter 2 of this thesis aimed at investigating whether indeed interlocutors initiate speech planning while still listening. The findings of this chapter provided clear evidence that the speed with which a turn can be taken over depends on when the response could start being prepared (planning onset time). These findings suggest that participants were indeed planning while listening, a finding that supports the recent Levison & Torreira model (2015). At this point it is important to note that the research for chapter 2 was conducted in 2011. At that time there was no solid empirical evidence published yet regarding this question. But many recent studies have since then confirmed this idea (Bögels et al., 2018, 2015; Boiteau et al., 2014; Sjerps & Meyer, 2015).

What factors impact on turn taking timing and especially the timing of response planning?

The unique contribution of chapter 2 to the question on whether planning initiates before the turn-end, is that it also revealed that onset of planning can interact with what I have labeled planning efficiency (assessed via a manipulation aimed at making conceptual retrieval more effortful, or by manipulating the syntactic complexity of the initial noun phrase to be produced). When initiating planning late, low planning efficiency affected turn-taking timing (see section below). But when planning was allowed to initiate early enough, any effects of low planning efficiency seemed to get absorbed by the extra time given to plan. As such, knowing that one has to prepare a demanding utterance might trigger early planning onset in order to achieve reasonable turn-taking timing.

The observed interaction between planning onset and efficiency is striking given the rather simple nature of the stimuli that were used. The materials in experiments 1 and 2 of chapter 2 consisted of only two sentence types (ones where information became available early or late), and within those types, the only difference between sentences was whether they indicated the left or right side pictures (and thus with rather constant effort on the listening side). I chose this rather basic sentence structure, because it allowed for a firm control of the planning onset time, thereby also allowing measurements of added effects of planning efficiency on turn-taking timing. Since these findings could already be observed with such basic and predictable structures it suggests to be a general property of combined speaking and listening. When looking into fragments of natural conversation, however, it is often hard to assess whether the registered gap duration represents simply a late planning onset or also effects of low planning efficiency (or even effects of high demands on the listening side).

Even though it has previously been demonstrated in language production research that planning initial complex noun phrases results in significantly longer onset latencies than planning initial simple noun phrases (Smith and Wheeldon ;1999), the current thesis (chapter 2) additionally demonstrates that the different demands linked to producing these syntactic structures may also interact with the planning of speech onset, when switching from listening only, to listening and preparing to speak. That is, if planning would start really early, even initial complex noun phrases might not take additional time to be articulated.

This being said, it is interesting to note that the processing outcome of initiating planning early seems to share some attributes with the processing outcome of so-called prepared speech in delayed naming tasks, in which participants have to buffer their response until prompted to respond. For example, when speech is already prepared (as opposed to online speech production), word frequency effects do not seem to impact on speech onset latencies or error rates (Jescheniak & Levelt, 1994). Likewise, when in chapter 2 (Experiment 2) speech was already prepared, because of early planning onset, no impact of conceptual retrieval difficulty was observed; while it was observed when planning onset was late.

At the same time, the processing outcome of initiating planning late (almost at turn-end), seems to share some attributes with the processing outcome of on-line speech. Prepared speech has been contrasted to online-speech, and it has been argued that these two types of speech invoke different processing strategies. For example, Wheeldon and Lahiri (1997) concluded that response time in prepared speech production depends on the number of phonological words a sentence contains, whereas response time in on-line speech production depends on the complexity of the to be produced phonological word. The idea that articulation of the formulated reply is buffered until the system recognizes that it is time to initiate articulation is also part of the Levinson-Torreira turn-taking timing model (Levinson & Torreira, 2015). Thus, depending on when in time the interlocutor could start planning, different planning strategies might emerge and result in online or prepared speech, which in turn might lead to different gap duration between turns.

A subtle manipulation of planning effort (and thus planning efficiency) was also included in chapters 3 and 4, even though these chapters focused on listening quality during concurrent planning and not on planning efficiency. Chapters 3 and 4 made use of high- and low- frequency picture names. This is a subtle manipulation of planning effort, as also evident in the observed robust effects on pupil dilation measures, with larger dilation for low- as compared to high-frequency picture-names. This subtle planning complexity manipulation affected response speed (my experimental approximation of turn-taking timing): low frequency picture-names were named slower than high frequency ones. It would be interesting to link this subtle planning effort manipulation to effects on the quality of listening (recognition memory performance). But due to the setup of the study, too few trials would be left to compare recognition memory for items heard in conjunction with naming high as opposed to low frequency pictures.

Future work should elaborate more on how and when efficiency of planning might affect turn-taking timing. For example, more demanding manipulations introducing more uncertainty regarding the syntactic structure to be produced, or manipulations in which a richer conceptual description is needed, could further test whether and when planning efficiency impacts on turn-taking timing. Note that the impact of more demanding planning efficiency manipulations could also be evaluated with regard to their direct impact on listening quality, rather than on turn-taking-timing; instead of looking at single word production as was the case in chapters 3, 4 and 5, the focus would now shift to production of more complex utterances. This is particularly important in order to understand the computational pressures exerted during conversational interaction.

Does listening performance drop, due to the need to prepare speech?

The demonstration that people indeed do plan, while still listening, seems to suggest that participants in conversation engage in a dual task situation, whereby they have to still listen, while already preparing a response. How do people perform in this dual task situation? Does listening performance drop, due to the need to prepare speech? This thesis clearly demonstrated that planning speech while listening to speech comes with a cost for listening quality, as in all cases items heard during planning were also less often recognized in the subsequent recognition memory test. This finding is in line with research suggesting that dual tasking impacts on consolidation or retrieval memory performance (e.g. Fernandes and Moscovitch, 2000) and research demonstrating that acoustic challenge can affect memory for heard words or syllables, even though these have been perceived correctly (e.g. Heinrich & Schneider, 2011; Heinrich, Schneider, & Craik, 2008; Surprenant, 1999).

Future work should focus more on memory of the comprehended content, rather than memory of the exact words used, since this is of essence in conversation. For example, showing an impact of concurrent planning on the comprehension or episodic memory of a scene description (e.g. who is doing what to whom), or demonstrating an even more pronounced Moses Illusion (see Erickson and Mattson, 1981 for details), would be an interesting next step. The term *Moses Illusion* has been used to describe the phenomenon by which people fail to identify a rather major discrepancy in something they are asked or told. For example, when asked "How many animals of each kind did Moses take on the ark?" people often fail to notice that it was actually Noah and not Moses that took the animals on the ark. Instead they respond to the part of the sentence that is requesting to give a number (see also Redner & Kusbit, 1991). An other interesting approach to evaluate the impact of planning on listening might be to focus on whether concurrent planning can make interlocutors miss out more often on a speech error by their conversational partner (see Ganushchak and Schiller, 2010 for detecting speech errors in the speech of others).

Can the observed impact of planning on listening performance be traced online/in real time?

The impact of planning on recognition memory performance (chapters 3 and 4) indicates that speech planning did to some degree interfere with listening. In chapter 5 I showed that the impact of concurrent planning on listening can also be registered online. A significant decrease in the amplitude of the N400 semantic anomaly-effect was registered for items heard during speech planning compared to items heard when not engaged in planning. The online nature of this finding makes it perhaps even more relevant to conversational research. At the same time the study of chapter 5 again confirmed an impact on recognition memory performance. This seems to suggest that the offline measure chosen to evaluate listening quality, which was also used in chapters 3 and 4, does nicely match the online measure; items heard during planning were processed less efficiently and this could be registered both online and offline.

An influence of planning speech on thorough semantic processing has also been reported in a very recent study by Bögels, Casillas, and Levinson (2018). Using the novel quiz game paradigm also used in Bögels and colleagues (2015) the authors manipulated planning onset timing (early vs. late) to investigate the impact of planning on comprehension processes. Their findings suggest that an impact of production planning on listening was mainly evident in so called quick responders, whose N400 size effect was smaller than that of slow responders, suggesting that the former shifted attention away from listening earlier than the latter. In chapter 5 I found that concurrent planning overall strongly reduced the N400 amplitude.

So why could this difference be registered for the majority of participants in chapter 5, but only in the subset of fast responders in the study by Bögels and colleagues (2018)? One possible explanation is that unlike the quiz paradigm, the study in chapter 5 might have made it easier for participants to focus on production rather than comprehension, even though they were instructed to also listen carefully, since they would later "have to do something" with these sentences. As a result, attention may have switched away earlier from listening. Note however, that recognition memory performance in the plan task of chapter 5 was still at an acceptable level (61.5%), which indicates that participants did not fully ignore the heard sentences. Moreover, recognition memory performance was still mediated by predictability, which implies that the sentence-final-words were nonetheless processed to some degree.

Another explanation may come from differences in the trial-structure of these studies. In chapter 5 the timepoint at which planning could start and the timepoint at which the predictable word was heard exactly coincided. This was not the case in the Bögels and colleagues (2018) study. In their study the point at which a word could be anticipated was either before or after planning could be initiated. Crucially, the absence of exact overlap between the point of perceiving the anticipated or unanticipated word, and perceiving the critical word that would trigger the planning response, might have allowed for the N400 effect to survive the early-planning condition (see the discussion section of chapter 5 for a more detailed description). It seems thus that capturing the impact of planning on listening online via the N400 manipulation is sensitive to the timing of the various components at work. Maybe exact overlap of planning onset and anticipated content is most detrimental, while subtle shifts in timing allow for more degrees of freedom in dealing with this dual-task situation. More research is needed to further elucidate this possibility.

Is there a tradeoff/interference between listening and planning to speak?

Manipulations of either the planning or the listening efficiency could reveal a trade-off between performance in listening and performance in speech production. That is, rather than affecting the timing of turn-taking, planning efficiency might also directly impact listening quality; and listening efficiency might affect production quality. Indeed, the experiments in chapter 3 demonstrated a trade-off in listening and production. That is, relative to Experiment 1 (incidental encoding), Experiment 2 (intentional encoding) resulted in better recognition memory (see figures 3 and 6, chapter 3). Importantly, however, this increase in listening performance was linked to poorer production performance (accuracy, see figure 8, chapter 3). Interestingly, response times in the exposure phase were not affected, which indicates that this trade-off cannot be explained by differences in response speed.

Another kind of trade-off has been observed in research on how dual tasking affects the N400 effect (e.g. Batterink et al., 2010; Giesbrecht et al., 2007; Hohlfeld et al., 2004; Lien et al., 2008; S. J. Luck et al., 1996; Rolke et al., 2001; Vachon & Jolicoeur, 2011; Vachon & Jolicœur, 2012) and suggests that the N400 component amplitude is reduced when attention is shared with a secondary task. A strong reduction in the N400 component amplitude was also seen in chapter 5 of this thesis. Importantly, the findings of chapter 5 are one of the first to show that the N400 amplitude can be so severely affected by the need to combine two linguistic tasks, rather than a linguistic and a non-linguistic task. Note that the tasks used in chapter 5, are not too far from what speakers are asked to do in every day conversation: listen and prepare to speak.

These observations are not only important as contributions to possible mechanisms implicated in conversational interaction. They are also important in as far as they point out that language processing in a setting that includes both listening and preparing to speak operates in more complex and more varied ways than when looking at listening or speaking in isolation, as has been the standard in most psycholinguistic research.

One of the few psycholinguistic settings in which speech production has been indirectly combined with speech perception (or visual word recognition), is the so called picture-word interference paradigm. This has been used to address questions regarding the order and timing of speech planning processes. Studies using the picture-word interference paradigm have shown that single word production can be affected by the simultaneous presentation of spoken words (e.g. Damian & Martin, 1999). For example, when picture presentation and auditory distractor word are semantically related and either overlap in time, or the distractor precedes picture presentation, then naming is slowed down (semantic interference); while when picture presentation and auditory distractor word are phonologically related and the distractor is presented in overlap or after the tagert, naming is faster (phonological facilitation; Schriefers et al., 1990). Chapters 3, 4 and 5 of this thesis demonstrated that not only picture names can be affected by heard words (as suggested by the picture-word interference paradigm), but also heard words can be affected from preparing a picture name. In particular, chapters 3 and 4 provided evidence that overall listening quality offline, as measured by recognition memory performance, was negatively affected by concurrent planning.

It seems thus that not only production is affected by simultaneous presentation of spoken words (mostly a domain-specific language effect), but that speaking and listening are tasks that compete for domain-general cognitive resources. As such, an impact can be registered not only from listening to speaking, but also the other way around. Future research might look into how the impact of speaking on listening is affected by semantic or phonological relatedness, thereby also offering more insight in domain-specific processes and on why and when a processing cost for listening emerges.

What is the source of the interference?

Based on the studies conducted in this thesis it is, unfortunately, not clear which components of the word recognition processes were most strongly affected by concurrent speech planning, and which aspects of the speech planning processes could have caused the disruption of the processing or storage of the spoken words in chapters, 3, 4 and 5.

Importantly, regarding the impact of concurrent planning on listening as registered in chapter 3, one could look for potential candidates of interference at the domain-specific level (e.g. early conceptual and/or subsequent lexical retrieval processes) or at the domain-general level (e.g. switching attention between tasks). Indications that the domain-specific level might contribute to the interference come from the aforementioned studies looking at how single word production is affected by the simultaneous presentation of spoken words (e.g. Damian & Martin, 1999; Schriefers, Meyer, & Levelt, 1990), but also from studies suggesting that the neuronal infrastructures of speaking and listening show substantial overlap at the semantic, lexical and syntactic level (Menenti et al., 2011). Indications that the domain-general level might contribute to interference come from 1) studies reporting that recognition memory is poorer when attention had been divided during encoding (see for example, Fernandes and Moscovitch, 2000); 2) research pointing to a role of attention in spoken word planning (see Roelofs and Piai, 2011, for a review); and 3) research reporting a reduction in the N400 component, when attentional resources are withdrawn from the semantic properties of words (Sommer and Hohlfeld, 2008; Hohlfeld and Sommer, 2015; Luck, 1998; see also Kutas and Federmeier, 2011; Van Petten, 2014 for a review of studies that focused on how divided attention reduces-or even eliminates-N400 effects).

Future research should aim at elucidating not only when, but also how the need to share resources between listening and planning results in interference between the two, and how it interacts with turn-taking timing. That would require a step-by-step build up of experiments aimed at looking into how, for example, components of the planning process (conceptualization, lexical selection, phonological encoding, phonetic encoding) have an impact on listening and vice versa, or at how domain-general aspects such as attention sharing or attention switching come into play.

Can predictability help listeners cope with those dual-task demands?

This thesis has demonstrated that planning speech while listening to speech results in interference. As a result the quality of listening decreases and even trade-offs between listening and planning can be observed. What possible ways are there to overcome this interference and the trade-offs resulting from this dual-task situation? One mechanism that has been linked to processing facilitation is predictability.

A number of studies have demonstrated that predictable words are processed faster than less predictable ones (see for example Altmann & Kamide, 1999; Kliegl, Nuthmann, & Engbert, 2006; Traxler & Foss, 2000). Moreover, recent research has described prediction as 1) supporting a timely planning initiation via assisting comprehension (Levinson & Torreira, 2015); or 2) directly supporting a timely planning initiation via content prediction (Corps et al., 2018), or 3) supporting a timely response initiation via content prediction, which allows for a successful turn-end projection (Magyari et al., 2014, 2017; Magyari & de Ruiter, 2008). Moreover, Ferreira and Lowder (2016) proposed a framework that brings together the notions of information structure, superficial (good-enough) language processing and prediction. They claim that prediction assists comprehension of the constituents with the new information.

To elaborate, in attempting to explain how interlocutors manage to take over a turn with minimal gap duration, Levinson and Torreira concluded that comprehenders use prediction to extract the speech act of the incoming utterance at the earliest possible point (Levinson & Torreira, 2015). Having extracted the speech act will then allow them to initiate speech planning, so that a sufficient amount of utterance is already ready to be articulated once the turn-final cues signaling the end of the incoming utterance are perceived; thus a fast turn-taking is accomplished.

On a similar note to Levinson and Torreira (2015), recent literature proposes that in conversation comprehenders allocate their information processing resources differently, depending on whether information is given or new, with given information only being processed at a superficial level (this has been termed the 'good-enough' approach, Ferreira & Lowder, 2016). The resources freed by this good-enough processing of the given information are then used to generate predictions that will facilitate comprehending the new information. Finally, content prediction has been proposed to contribute to faster turn-taking in two ways; by making it possible to initiate planning earlier (Corps et al., 2018), or by allowing for an accurate turn-end projection and thus a timely response initiation (Magyari et al., 2014, 2017; Magyari & de Ruiter, 2008).

In this thesis, the potential impact of predictability on planning initiation and response initiation, as well as its role in comprehending new information, was only indirectly addressed. The focus was on how prediction might facilitate listening, and thus make it less prone to interference from speech planning (chapters 4 and 5). I decided to use the term "predictability" here, but it is important to note that it is not easy to conclude whether any effects observed should be directly ascribed to *prediction* or rather to *integration* difficulties (Kutas, Delong, & Smith, 2011).

In chapter 4 a rather subtle predictability manipulation was introduced, which compared predictable to unpredictable, but still plausible, sentence-final-words. No impact of this kind of predictability on recognition memory was registered. Thus, this subtle predictability manipulation does seem not to have rendered listening much easier compared to the unpredictable but plausible sentence-final-words. As a result, listening quality was not improved to a degree that would influence memory encoding. Even so, predictability affected response speed (turn-taking timing in a sense), since naming a low frequency picture was faster when having just heard a predictable sentence-final-word compared to when having just heard an unpredictable sentence-final-word. The fact that the predictability effect was not seen overall for speech planning of low and of high frequency picture-names, but only for low frequency picture-names, suggests that facilitation due to predictability might at times only be noticeable when the speech planning part is more demanding (low frequency, as opposed to high frequency picture-names in this case). So even if the facilitation is not directly noticeable as a memory gain on the listening side, it can still be traced on the planning side, when speech planning is most demanding (low frequency picture-names). This possible facilitation in listening allowed for more resources to be put into speech planning, when planning got most demanding (low frequency picture-names). Thus, a direct benefit could not be registered on the listening quality side, but was registered as a carry-over effect on the speech planning side, by allowing for more resources to be put into planning low frequency picture-names.

In chapter 5 a more drastic distinction was implemented, as expected sentence-final-words were compared to semantically anomalous ones, which were not plausible in any way as closings of these sentences. This kind of manipulation did affect listening efficiency to a degree that an impact on recognition memory performance was seen: Among the subset of sentence-final-words that had been presented on picture naming trials, those words that were expected given their preceding context were better remembered than the semantically anomalous ones. At the same time, response speed (in a sense, turn-taking timing) was also affected (exposure phase), as listening to expected sentence-final-words led to faster picture naming than listening to anomalous sentence-final-words. Thus, unlike chapter 4, in chapter 5 a benefit was registered not only on the planning but also on the listening quality side (recognition memory performance). It seems thus that the type of predictability manipulation introduced plays a role in whether it will contribute to facilitate listening and on how any resulting heightened listening efficiency will impact on other processes (e.g. speech planning). Note that when comparing expected to semantically anomalous sentence-final-words, successful turn-end projection might also play a role; in the expected sentence-final-word one has optimal content prediction and thus also optimal turn-end anticipation (following Magyari et al., 2014, 2017; Magyari & de Ruiter, 2008), while in the semantically anomalous sentence-final-word the predicted content is overwritten by having to adapt to a semantically anomalous input and the turn-end projection is not borne out.

The finding of chapter 5 that predictable sentence-final-words did speed up speech planning of completely unrelated picture-names and of chapter 4 that predictable sentence-final-words sped up speech planning of low frequency picture-names seems to fit the role ascribed to prediction by the Levinson - Torreira turn-taking model . When listening is made easier through predictability, then speech preparation can initiate at an earlier timepoint. Even though our material was not extracted from conversation corpora and the auditory input was not contingent to the to be produced picture name, still a benefit from predictability was registered. Importantly, due to the experimental design and stimuli chosen, this benefit can be solely attributed to predictability and is not confounded with other factors that could affect the timing of turn-taking.

Importantly, the research in chapters 4 and 5 suggests that prediction -at least when the moment of planning initiation and of listening to the predictable word completely overlap- is not only a tool to speed up comprehension and thus initiate production planning as early as possible (Corps et al., 2018; Levinson & Torreira, 2015), and not only a tool to facilitate comprehending new information (Ferreira & Lowder, 2016). Prediction also contributes to reducing the interference between listening and speech planning, as predictable words were remembered better than unpredictable in chapter 5 and picture-names could be named faster when the sentence-final-word was predictable (chapter 5; chapter 4, only for low frequency picture-names).

Combining the findings of chapters 4 and 5 with the ideas proposed by Levinson and Torreira (2015), Magyari and de Ruiter (2008), Corps and colleagues (2018) and Ferreira and Lowder (2016), lead me to propose the following pattern of interactions: good-enough processing of the given information in the sentence frees resources for predicting, which will not only facilitate comprehending new information (Ferreira and Lowder, 2016), but also initiating concurrent planning (Corps et al., 2018; Levinson & Torreira, 2015), possibly taking over the turn more timely (Magyari et al., 2014, 2017; Magyari & de Ruiter, 2008); and all that with minor impact on the heard input (low interference between listening and planning).

Findings like the ones of chapter 4, in which producing low frequency picture-names was faster when listening to predictable sentence-final-words- and from chapter 5, in which listening to expected sentence-final-words allowed for faster naming overall, indicate that not only planning efficiency manipulations but also listening efficiency can affect the timing of turn-taking. Thus, more efficient listening due to listening facilitation (predictable words) affected response speed (my experimental approximation of turn-taking timing) to a picture-name that was totally unrelated to that predictable-heard-word. Many previous studies have shown that processing predictable words is faster than processing less predictable ones (see for example Altmann & Kamide, 1999; Kliegl, Nuthmann, & Engbert, 2006; Traxler &

Foss, 2000). The current research, however, demonstrates that the processing advantage of predictable words can transfer from the listening side to the production side when listening and production planning are combined in one experimental paradigm.

If, then, one can think of predictability as a way to make listening more efficient, an interesting future research topic would be to identify other ways in which listening efficiency can be manipulated. For example, more uncertainty regarding the syntactic structure to be comprehended, or adding external environmental noise as is mostly the case in natural conversation, could further test whether and when listening efficiency influences the timing of turn-taking.

Monitoring the cognitive effort of planning while listening, via pupil dilation. The main aim of this thesis was to understand the processing constraints when having to prepare speech while still listening. In studying this question, chapters 3 and 4 relied on pupillometry to monitor the cognitive effort of planning while listening, an approach that is relatively novel to the field of turn-taking behavior. It was found that planning speech while listening induced larger pupil dilation than just listening. Moreover, it was found that planning low-frequency picture-names resulted in larger pupil dilation than planning high-frequency picture-names. Experiment 2 of chapter 4 confirmed that the observed pupil modulations were indeed the product of planning effort and not of some lower level processes like picture recognition. The fact that pupil size modulation by planning effort (with planning trials resulting in larger pupil dilation than no-planning trials) matched modulations of recognition memory performance by planning effort (with planning items recognized less often than no-planning items) offered some evidence that indeed planning effort is one important factor influencing listening quality (evaluated via recognition memory performance). This holds in particular for the need to effectively share general cognitive resources between planning and listening.

Task Evoked Pupillary Responses (TEPR's) have at times been used in psycholinguistic research (e.g. Engelhardt, Ferreira, & Patsenko, 2010; Hyönä, Tommola, & Alaja, 1995; Kuchinke, Võ, Hofmann, & Jacobs, 2007; Kuipers & Thierry, 2013; Papesh & Goldinger, 2012; Tromp et al., 2016) but the current work demonstrates that they might also prove to be a useful tool in investigating cognitive effort during linguistic dual-task situations, such as turn-taking in conversation.

Recently pupillometry has also been used to monitor listening effort in challenging listening conditions, such as when different kinds of noise are overlapping speech, or when individuals have hearing impairment (see Peelle, 2018 and Winn & Moore, 2018 for an overview). The "framework for understanding effortful listening (FUEL), defines listening effort as the purposeful resource allocation of mental capacity to a given listening task, when encountering adverse listening conditions (Pichora-fuller et al., 2016). This thesis is the first to consider that listening effort can also be affected by the need to allocate resources to speech planning. As such, listening to conversational speech under challenging listening conditions (elevated noise, multi-talker environment or even hearing impairment) should be even more effortful, because one not only has to allocate resources to listening, but should also produce a timely response. Even though having to cope with challenging listening conditions while engaging in conversation is a rather common situation, future research should look into how listening effort is modulated by the need to shift resources to speech preparation. Such research might also help in pinpointing the components that contribute most to elevating listening effort in hearing-impaired individuals, when participating in conversation. According to Wierda, Rijn, Taatgen, and Martens (2012) pupil dilation can be used to track attention at high temporal resolution. Thus, one could possibly use it to establish how cognitive load is distributed during listening and speaking in a dialogue situation.

Limitations and merits of the applied paradigm.

The research reported in this thesis has been based on carefully designed experiments that created quite constrained settings within which participants planned speech while listening. I used this strict design to be able to measure subtle effects of processing constraints on 'turn-taking' behavior. However, as a result, these experiments lack the interactional contingency which is an eminent characteristic of natural conversation. Recently, research on conversation is increasingly addressing similar questions in more natural experimental designs (for example, the quiz game paradigm used by Bögels et al., 2018, 2015). This could be important because everyday conversation adheres to interactional constraints. For example, non-compliance to a request can be signaled by the interlocutor by leaving a long pause before replying (Roberts, Torreira, and Levinson, 2015). Since people are presumably aware of this potential interpretation, this provides a strong incentive to start planning a reply soon (and, hence, mimimize the gap between turns). Thus, turn-taking-timing in real conversation (including the conversations taken up in conversational corpora) is not only the product of processing constraints but also of interactional constraints. The increased use of such designs thereby provide a promising avenue to investigate some of the important aspects of turn-taking that could not be measured in the carefully controlled but more artificial turn-taking setup used here.

In addition to the signaling function of pauses, there are also other characteristics, which may affect turn-taking behavior by providing facilitation to the turn-taking dynamic. One such characteristic is common ground. Common ground refers to all the information that interlocutors believe they share with each other (Clark, 1996) and which is meant to facilitate communication (Clark and Brennan, 1991), as it allows anticipating what the conversational partner knows and what not. Nonetheless, it has been argued that since retaining common ground requires that interlocutors explicitly model each other's beliefs, this is in itself a resource consuming process (Garrod and Pickering, 2004), which will probably not only facilitate conversational interaction, but will at times take away resources from other processes like listening and preparing a reply.

Another characteristic of natural conversation that was not implemented here is information structure. As was already discussed in the section on the role of predictability, sentences in discourse typically include a given and a new part. The given part is part of what has already been established and referred to in the course of the discourse, while the new part is the new information entering the discourse (Ferreira & Lowder, 2016). Information structure can be conveyed in a number of ways, amongst which is prosody (e.g. Wang, Bastiaansen, Yang, & Hagoort, 2011) and word positioning in a sentence. For example, there is a tendency to place given information at the beginning of a sentence and new information at the end. In fact, as already pointed out, Ferreira and Lowder (2016) argue that given information is processed in good enough manner in order for resources to be spared for predicting the new part of the sentence. Thus, information structure may allow for an efficient resource distribution, hence facilitating conversation. For example, the resources spared by processing the given information in a good enough only manner, could at times be spent in predicting the new part of the sentence, but also in preparing speech. That could mean that interlocutors might more often plan during the given part of the sentence, than during the new part of the sentence.

Finally, the interactive alignment model of dialogue (Garrod & Pickering, 2004) argues that conversation is easy, despite the fact that the interlocutor is faced with a number of tasks. It argues that interlocutors are in fact participants in a joint activity (see also Clark, 1996). As such, they automatically align their situation models, which allows for a smooth conversational interaction. It is through the interactive nature of dialogue that linguistic representations are interactively aligned. And these aligned representations are more easily computed than non-aligned representations, because they are computed by reusing information already computed by the other interlocutor. As a result the processing load is distributed between the interlocutors. The alignment can happen at a number of levels, ranging from phonology, to syntax and semantics and results in making comprehension and production in conversation easier (Garrod & Pickering, 2004). One of the central concepts of interactional alignment is priming. For example, interlocutors tend to re-use specific expressions, as the conversation unfolds. These expressions retain this 'routinized' character only for the particular interaction and are thought to make both production and comprehension easier. Actually the authors hypothesize that interlocutors might be able to partly skip some parts of the production processes in the same way as they can do that for comprehension (Pickering & Garrod, 2004), thereby rendering production and comprehension within conversation easier. Again, this aspect, which is presumably most predominant in natural, everyday conversation was not assessed in this thesis, but could help alleviate some of the processing constraints that were outlined.

On the other hand, there are aspects of everyday conversation which were not part of this study, but which would most probably render the conversational interaction more effortful. Processing spoken language in actual conversational environments requires coping with all kinds of added noise, such as background noise, listener and speaker characteristics, which often result in speech signal degradation (Strand, Brown, Merchant, Brown, & Smith, 2018; for a review on speech recognition in adverse conditions see Mattys, Davis, Bradlow, & Scott, 2012). This aspect implies that processing spoken language in conversational environments can actually be even more cognitively challenging than what was assessed in this thesis. So the cost of having to plan while still listening under these varying conditions might be even bigger. In fact it might even dictate whether planning while listening is possible or not.

One might argue that all the characteristics described above, are so intricately linked to natural conversation that leaving them out does not allow for solid conclusions. So, what is there to gain from studies that leave out important aspects of the conversational setting, like interactional contingency, common ground, information structure, alignment aspects and even speech signal degradation? I argue that conducting controlled psycholinguistic experiments, which leave out the interaction constraints and other aspects of the conversational setting, but combine listening and production planning of controlled stimuli, can still offer important insights in possible processing constraints affecting turn-taking timing in natural conversation. For example, by having complete overlap of listening and planning, it was possible to rule out any effects that could stem from buffering a response until a turn-end is detected; as buffering a response might itself be a resource-consuming enterprise that could affect both planning and listening. In addition, by controlling predictability of our stimuli it was possible to evaluate the contribution of this factor independently of other possible sources of facilitation or interference.

Moreover, alignment processes might be argued to be mostly automatic but there is also a strategic component involved (Garrod & Pickering, 2007). As such these are not completely resource free processes and also have an impact on cognitive effort. Thus, if one was to include them it would be hard to control if and when they contribute to adding or subtracting cognitive effort, when listening and preparing to speak at the same time. As such, the findings of this thesis might not be applicable directly to natural conversation, but can point to specific factors and mechanisms that should then be investigated in a more interactional setting that would have to be built up step by step. In line with this, Levinson and Torreira (2015) have argued that any psychologically realistic model of turn-taking is bound by the constraints imposed by psycholinguistic processing. In fact, a number of findings reported in recent research using the quiz game paradigm (Bögels et al., 2018, 2015) seem to be in line with the findings of this thesis.

Finally, it would be interesting to consider the possibility that processing costs due to the need to combine planning and listening (like those observed in this thesis), could also be one of the driving forces in computing aligned representations, in order to reduce computational load. Future research should investigate whether the degree or level of alignment between interlocutors is linked to such computational pressures. These computational pressures might also partly explain why conversation is highly repetitive (e.g. Tannen, 2007) and why phenomena like lexical entrainment (Brennan, 1996) and syntactic persistence occur (Bock and Griffin, 2000; Bock, 1986).

Implications for the conversational setting. The interlocutor as a multitasker. As stated in chapter 1, the focus of this thesis was on the interface between listening and concurrent planning. As such, the materials and paradigms used were more focused on offering the opportunity to independently manipulate a number of individual factors, and much less on including parameters that would make them more conversation-like. Yet, the questions addressed in this thesis capture part of the processing constraints that emerge in conversation and as such provide important links to basic mechanisms involved in successful conversational interaction.

Overall the findings of this thesis demonstrate that even though interlocutors perform quite well as multi-taskers -effectively switching between listening and preparing to speak- some costs can be traced, for both listening and for planning of speech. These costs might not be directly noticeable as such in conversation (conversation is usually experienced as smooth and easy), but are nonetheless factors that constrain conversation behavior; decisions on when to initiate planning, on how much to focus on listening or on production planning, and on the complexity of the response. That is, if the interlocutor wishes to cope with the demands introduced by this dual situation, she needs to actively adapt her behavior. For example, when the incoming signal (listening) is easy to follow, fast preparation of one's turn might be feasible. But when the incoming signal (listening) becomes more complex, shifting more resources onto listening might be important, hence decreasing the amount of resources available for early production planning.

The idea of prioritizing one process over the other has also been argued to play a role in the need to share central resources between conversing and driving (Becic et al., 2010). Moreover, the cost of having to switch attention towards one or the other process has also been addressed by Vachon and Jolicoeur (2011, 2012), who postulate that in dual tasking situations semantic processing is affected by task switching during which the task set needs to be reconfigured. When this has to happen fast (due to short SOA) the impact on semantic processing is even stronger. Thus, at times, the specific goals pursued by the interlocutor might make her focus less on the structural difficulties introduced by the incoming signal (e.g. complex linguistic input), and instead prioritize fast turn-taking.; for example, if one wishes to retain the floor. Note however that this might come at a cost of listening quality.

It is, hence, most likely that most of the times an interlocutor cannot avoid having a cost in one or more processing outcomes. These costs also raise an important point concerning current psycholinguistic theories. If the ultimate goal is to understand how language operates in face-to-face communication (Holler et al., 2016), then a crucial first step would be to investigate listening and speaking in common paradigms, rather than only in isolation. One psycholinguistic theory that has addressed the issue of capacity limitations of the comprehension system and partly also of the production system is the “good enough” approach to language processing, which has recently also been linked to the notions of information structure and prediction (Ferreira & Lowder, 2016). According to this approach the end product of sentence comprehension processes is tuned to the addressees' particular goals. As such the capacity limitations of the comprehension system might trigger the formation of a strategically underspecified sentence-representation, which, importantly, is “good enough” for pursuing one's current goals. So the language-processing system adjusts in a way such that its limited resources are distributed according to the current goals (see Ferreira, Bailey, & Ferraro, 2002; Ferreira & Patson, 2007 for a discussion on good enough perception, and Swets, Jacovina, & Gerrig, 2012 for discussion on good enough production). Hence, most of the time, the costs observed in either listening, production, or in the timing of turn-taking, will be of a kind that the interlocutor can tolerate for the sake of smooth conversation.

References

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs : restricting the domain of subsequent reference. *Cognition, 73*, 247–264. https://doi.org/10.1016/S0010-0277(99)00059-1

Baayen, R.H., Piepenbrock, R., van Rijn, H. (1993). The CELEX Lexical Database [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Baddeley, A. (1976). *The psychology of memory*. New York. Harper & Row.

Baddeley, A., Lewis, V., Eldridge, M., & Thomson, N. (1984). Attention and Retrieval From Long-Term Memory. *Journal of Experimental Psychology. General*, *113*(4), 518–540.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barthel, M., Meyer, A. S., & Levinson, S. C. (2017). Next Speakers Plan Their Turn Early and Speak after Turn-Final " Go-Signals ." *Frontiers in Psychology*, *8*(393). https://doi.org/10.3389/fpsyg.2017.00393

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., … Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, *10*(2), 344–380. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12921412

Batterink, L., Karns, C. M., Yamada, Y., & Neville, H. (2010). The Role of Awareness in Semantic and Syntactic Processing : An ERP Attentional Blink Study. *Journal of Cognitive Neuroscience*, *22*(11), 2514–2529. https://doi.org/10.1162/jocn.2009.21361

Beattie, G., Cutler, A., & Pearson, M. (1982). Why is Mrs Thatcher interrupted so often? *Nature*, *300*, 744–747. https://doi.org/10.1038/300744a0

Beattie, G. W., & Barnard, P. J. (1979). The temporal structure of natural telephone conversations (directory enquiry calls). *Linguistics*, *17*(3–4), 213–229. https://doi.org/10.1515/ling.1979.17.3-4.213

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 142–162). New York: Cambridge University Press. Retrieved from http://psycnet.apa.org/psycinfo/2000-03927-005

Becic, E., Dell, G. S., Bock, K., Garnsey, S. M., Kubose, T., & Kramer, A. F. (2010). Driving impairs talking. *Psychonomic Bulletin & Review*, *17*(1), 15–21. https://doi.org/10.3758/PBR.17.1.15

Bentin, S., Kutas, M., & Hillyard, S. (1995). Semantic Processing and Memory for Attended and Unattended Words in Dichotic Listening: Behavioral and Electrophysiological Evidence. *Journal of Experimental Psychology. Human Perception and Performance*, *21*(1), 54–67. https://doi.org/10.1037//0096-

1523.21.1.54

Besson, M., Kutas, M., & Van Petten, C. (1992). An Event-Related Potential (ERP) Analysis of Semantic Congruity and Repetition Effects in Sentences. *Journal of Cognitive Neuroscience*, *4*(2), 132–149. https://doi.org/10.1162/jocn.1992.4.2.132

Bock, K., & Levelt, W. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 945–984). London: Academic Press.

Bock, Kathryn, Griffin, Z. M. (2000). The persistence of structural priming: transient activation or implicit learning? *Journal of Experimental Psychology*, *129*(2), 177–192. https://doi.org/10.1037//0096-3445.129.2.177

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*(3), 355–387. https://doi.org/10.1016/0010-0285(86)90004-6

Bock, K., Dell, G., Garnsey, S., Kramer, A. F., & Kubose, T. T. (2007). Car talk, car listen. In A. S. Meyer, L. R. Wheeldon, & A. Krott (Eds.), *Automaticity and Control in Language Processing* (pp. 21–42). Psychology Press.

Boersma, P, Weenink, D. (2012). Praat: doing phonetics by computer. Retrieved from http://www.praat.org/

Bögels, S., Casillas, M., & Levinson, S. C. (2018). Planning versus comprehension in turn-taking : Fast responders show reduced anticipatory processing of the question. *Neuropsychologia*, *109*, 295–310. https://doi.org/10.1016/j.neuropsychologia.2017.12.028

Bögels, S., Magyari, L., & Levinson, S. C. (2015). Neural signatures of response planning occur midway through an incoming question in conversation. *Scientific Reports*, *5*, 12881. https://doi.org/10.1038/srep12881

Bögels, S., & Torreira, F. (2015). Listeners use intonational phrase boundaries to project turn ends in spoken interaction. *Journal of Phonetics*, *52*, 46–57. https://doi.org/10.1016/j.wocn.2015.04.004

Boiteau, T. W., Malone, P. S., Peters, S. a, & Almor, A. (2014). Interference between conversation and a concurrent visuomotor task. *Journal of Experimental Psychology. General*, *143*(1), 295–311. https://doi.org/10.1037/a0031858

Bosch, L. Ten, Oostdijk, N., & Boves, L. (2005). On temporal aspects of turn taking in conversational dialogues. *Speech Communication*, *47*(1–2), 80–86. https://doi.org/10.1016/j.specom.2005.05.009

Bosker, H. R., Reinisch, E., & Sjerps, M. J. (2017). Cognitive load makes speech sound fast , but does not modulate acoustic context effects. *Journal of Memory and Language*, *94*, 166–176. https://doi.org/10.1016/j.jml.2016.12.002

Brennan, S. (1996). Lexical entrainment in spontaneous dialog. In *International Symposium on Spoken Dialogue (ISSD-96)* (pp. 41–44).

Chwilla, D. J., Brown, C. M., & Hagoort, P. (1995). The N400 as a function of the level of processing. *Psychophysiology*, *32*, 274–285. https://psycnet.apa.org/doi/10.1111/j.1469-8986.1995.tb02956.x

Clark, H H, Brennan, S. E. (1991). Grounding in communication. In *Perspectives on socially shared cognition* (pp. 127–149). Washington, DC, US: American Psychological Association.

Clark, H. H. (1996). Using language. Cambridge, England: Cambridge University Press.

Cleland, A. A., Tamminen, J., Quinlan, P. T., & Gaskell, M. G. (2012). Spoken word processing creates a lexical bottleneck. *Language and Cognitive Processes*, *27*(4), 572–593. https://doi.org/10.1080/01690965.2011.564942

Connolly, J. F., Stewart, S. ., & Phillips, N. (1990). The Effects of Processing Requirements on Neurophysiological Responses to Spoken Sentences. *Brain and Language*, *39*, 302–318. https://doi.org/10.1016/0093-934x(90)90016-a

Cook, A. E., & Meyer, A. S. (2008). Capacity demands of phoneme selection in word production: new evidence from dual-task experiments. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *34*(4), 886–899. https://doi.org/10.1037/0278-7393.34.4.886

Corley, M., MacGregor, L. J., & Donaldson, D. I. (2007). It's the way that you, er, say it: hesitations in speech affect language comprehension. *Cognition*, *105*(3), 658–668. https://doi.org/10.1016/j.cognition.2006.10.010

Corps, R. E., Crossley, A., Gambi, C., & Pickering, M. J. (2018). Early preparation during turn-taking : Listeners use content predictions to determine what to say but not when to say it, *175*(January), 77–95. https://doi.org/10.1016/j.cognition.2018.01.015

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*(4), 491–511.

Couper-Kuhlen, E., & Selting, M. (1996). Towards an interactional perspective on prosody and a prosodic perspective on interaction. In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Conversation* (pp. 11–56). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511597862

Cutler, A., & Pearson, M. (1986). On the analysis of prosodic turn-taking cues. *Intonation in Discourse*, 139–155. Retrieved from http://pubman.mpdl.mpg.de/pubman/item/escidoc:76883:7/component/escidoc:506929/Cutler_1985_On the analysis.pdf

Damian, M. F., & Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *25*(2), 345–361. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10093206

de Ruiter, J. P., Mitterer, H., & Enfield, N. J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language*, *82*(3), 515–535. https://doi.org/10.1353/lan.2006.0130

Delong, K. a, Urbach, T. P., Groppe, D. M., & Kutas, M. (2011). Overlapping dual ERP responses to low cloze probability sentence continuations. *Psychophysiology*, *48*(9), 1203–1207. https://doi.org/10.1111/j.1469-8986.2011.01199.x

Duncan, S. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, *23*(2), 283–292. Retrieved from http://www.citeulike.org/user/laurap/article/4401193

Duncan, S., & Fiske, D. W. (1977). *Face-to-face Interaction: Research, Methods and*

Theory. John Wiley & Sons Inc. Retrieved from http://www.amazon.co.uk/Face-face-Interaction-Research-Methods/dp/0470991135

Dux, P. E., & Marois, R. (2009). The attentional blink: a review of data and theory. *Attention, Perception & Psychophysics*, *71*(8), 1683–1700. https://doi.org/10.3758/APP.71.8.1683

Einhäuser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(5), 1704–1709. https://doi.org/10.1073/pnas.0707727105

Engelhardt, P. E., Ferreira, F., & Patsenko, E. G. (2010). Pupillometry reveals processing load during spoken language comprehension. *Quarterly Journal of Experimental Psychology*, *63*(4), 639–645. https://doi.org/10.1080/17470210903469864

Erickson, T., Mattson, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior*, *20*(5), 540–551. https://doi.org/10.1016/S0022-5371(81)90165-1

Fernandes, M. A., & Moscovitch, M. (2000). Divided attention and memory: evidence of substantial interference effects at retrieval and encoding. *Journal of Experimental Psychology. General*, *129*(2), 155–176. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10868332

Ferreira, F., & Bailey, K. G. D. (2004). Disfluencies and human language comprehension. *Trends in Cognitive Sciences*, *8*(5), 231–237. https://doi.org/10.1016/j.tics.2004.03.011

Ferreira, F., Bailey, K. G. D. D., & Ferraro, V. (2002). Good-Enough Representations in Language Comprehension. *Current Directions in Psychological Science*, *11*(1), 11–15. https://doi.org/10.1111/1467-8721.00158

Ferreira, F., & Lowder, M. W. (2016). *Prediction , Information Structure , and Good-Enough Language Processing*. *The Pschology of Learning and Motivation* (Vol. 65). Elsevier Ltd. https://doi.org/10.1016/bs.plm.2016.04.002

Ferreira, F., & Patson, N. D. (2007). The "Good Enough" Approach to Language Comprehension. *Language and Linguistics Compass*, *1*(1–2), 71–83. https://doi.org/10.1111/j.1749-818X.2007.00007.x

Ferreira, F., & Swets, B. (2002). How Incremental Is Language Production? Evidence from the Production of Utterances Requiring the Computation of Arithmetic Sums. *Journal of Memory and Language*, *46*(1), 57–84. https://doi.org/10.1006/jmla.2001.2797

Ferreira, V. S., & Pashler, H. (2002). Central bottleneck influences on the processing stages of word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1187–1199. https://doi.org/10.1037//0278-7393.28.6.1187

Ford, C., & Thompson, S. (1996). Interactional units in conversation: Syntactic, intonational, and pragmatic resources for the management of turns. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and grammar* (pp. 135–84). Cambridge University Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=4m1qBJX1UJIC&oi=fnd&pg=PA134&dq=Interactional+units+in+conversation:+Syntactic,+intonational,+and+p

ragmatic+resources+for+the+management+of+turns&ots=QNRGGu2eZf&sig=jPcbjiNlIJfzQ9SjBCbuba2fjOE

Gambi, C., & Pickering, M. J. (2011). A cognitive architecture for the coordination of utterances. *Frontiers in Psychology*, *2*(November), 275. https://doi.org/10.3389/fpsyg.2011.00275

Ganushchak, L. Y., & Schiller, N. O. (2010). Detection of speech errors in the speech of others: an ERP study. *NeuroImage*, *49*(4), 3331–3337. https://doi.org/10.1016/j.neuroimage.2009.11.063

Garrod, S. & Pickering, M. (2007). Alignment in dialogue. Oxford Handbook of Psycholinguistics. 10.1093/oxfordhb/9780198568971.013.0026

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*(1), 8–11. https://doi.org/10.1016/j.tics.2003.10.016

Giesbrecht, B., Sy, J. L., & Elliott, J. C. (2007). Electrophysiological Evidence for Both Perceptual and Postperceptual Selection during the Attentional Blink. *Journal of Cognitive Neuroscience*, *19*(12), 2005–2018. https://doi.org/10.1162/jocn.2007.19.12.2005

Goldinger, S. D., Azuma, T., Abramson, M., & Jain, P. (1997). Open Wide and Say "Blah!" Attentional Dynamics of Delayed Naming. *Journal of Memory and Language*, *37*(2), 190–216. https://doi.org/10.1006/jmla.1997.2518

Goldinger, S. D., & Papesh, M. H. (2012). Pupil Dilation Reflects the Creation and Retrieval of Memories. *Current Directions in Psychological Science*, *21*(2), 90–95. https://doi.org/10.1177/0963721412436811

Gordon, P., Eberhardt, J., & Rueckl, J. (1993). Attentional modulation of the phonetic significance of acoustic cues. *Cognitive Psychology*, *25*, 1–42. Retrieved from http://www.sciencedirect.com/science/article/pii/S0010028583710017

Green, D., & Swets, J. (1966). *Signal detection theory and psychophysics* (Vol. 20). New York: Wiley.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, *304*(5669), 438. DOI: 10.1126/science.1095455

Heil, M., & Rolke, B. (2004). Unattended Distractor-Induced Priming in a Visual Selective Attention Task. *Journal of Psychophysiology*, *18*, 164–169. https://doi.org/10.1027/0269-8803.18.4.164

Heinrich, A., & Schneider, B. A. (2011). Elucidating the effects of ageing on remembering perceptually distorted word pairs. *Quarterly Journal of Experimental Psychology*, *64*(1), 186–205. https://doi.org/10.1080/17470218.2010.492621

Heinrich, A., Schneider, B. A., & Craik, F. I. M. (2008). Investigating the influence of continuous babble on auditory short-term memory performance. *Quarterly Journal of Experimental Psychology*, *61*(5), 735–751. https://doi.org/10.1080/17470210701402372

Hillyard, S. A., & Munte, T. F. (1984). Selective attention to color and location : An analysis with event-related brain potentials. *Perception & Psychophysics*, *36*(2), 185–198. https://doi.org/10.3758/bf03202679

Hintz, F., Meyer, A. S., & Huettig, F. (2016). Encouraging prediction during production facilitates subsequent comprehension : Evidence from interleaved object naming in sentence context and sentence reading. *Quarterly Journal of Experimental Psychology*, *69*(6), 1056–1063. https://doi.org/10.1080/17470218.2015.1131309

Hintz, F., Meyer, A. S., & Huettig, F. (2017). Predictors of Verb-Mediated Anticipatory Eye Movements in the Visual World. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(9), 1352–1374.

Hohlfeld, A., Sangals, J., & Sommer, W. (2004). Effects of Additional Tasks on Language Perception : An Event-Related Brain Potential Investigation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *30*(5), 1012–1025. https://doi.org/10.1037/0278-7393.30.5.1012

Hohlfeld, A., & Sommer, W. (2005). Semantic processing of unattended meaning is modulated by additional task load: evidence from electrophysiology. *Brain Research. Cognitive Brain Research*, *24*(3), 500–512. https://doi.org/10.1016/j.cogbrainres.2005.03.001

Hohlfeld, A., & Sommer, W. (2015). Is Semantic Processing During Sentence Reading Autonomous or Controlled ? Evidence from the N400 Component in a Dual Task Paradigm. *Advances in Cognitive Psychology*, *11*, 42–55.

Holler, J., Kendrick, K. H., Casillas, M., & Levinson, S. C. (Eds.). (2016). *Turn-taking in human communicative interaction*. Frontiers in Psychology. https://doi.org/10.3389/978-2-88919-825-2

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, *1626*, 118–135. https://doi.org/10.1016/j.brainres.2015.02.014

Hyönä, J., Tommola, J., & Alaja, A. (1995). Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *The Quarterly Journal of Experimental Psychology*, *48A*(3), 598–612. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/14640749508401407

Inaba, M., Nomura, M., & Ohira, H. (2005). Neural evidence of effects of emotional valence on word recognition. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *57*(3), 165–173. https://doi.org/10.1016/j.ijpsycho.2005.01.002

Indefrey, P. (2011). The spatial and temporal signatures of word production components: a critical update. *Frontiers in Psychology*, *2*(October), 255. https://doi.org/10.3389/fpsyg.2011.00255

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1–2), 101–144. https://doi.org/10.1016/j.cognition.2002.06.001

Jaeger, T. F. (2008). Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language*, *59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007

Janssen, N., Schirm, W., Mahon, B. Z., & Caramazza, A. (2008). Semantic interference in a delayed naming task: evidence for the response exclusion hypothesis. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *34*(1), 249–256. https://doi.org/10.1037/0278-7393.34.1.249

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 824–843. https://doi.org/10.1037//0278-7393.20.4.824

Just, M. A., Carpenter, P. a, & Keller, T. a. (1996). The capacity theory of comprehension: new frontiers of evidence and arguments. *Psychological Review*, *103*(4), 773–780. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8888654

Kafkas, A., & Montaldi, D. (2011). Recognition memory strength is predicted by pupillary responses at encoding while fixation patterns distinguish recollection from familiarity. *Quarterly Journal of Experimental Psychology (2006)*, *64*(10), 1971–1989. https://doi.org/10.1080/17470218.2011.588335

Kafkas, A., & Montaldi, D. (2012). Familiarity and recollection produce distinct eye movement, pupil and medial temporal lobe responses when memory strength is matched. *Neuropsychologia*, *50*(13), 3080–3093. https://doi.org/10.1016/j.neuropsychologia.2012.08.001

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, 1583–1585. http://doi.apa.org/psycinfo/1967-02420-001

Kellenbach, M. L., & Michie, P. T. (1996). Modulation of Event-Related Potentials by Semantic Priming : Effects of Color-Cued Selective Attention. *Journal of Cognitive Neuroscience*, *8*, 155–173.

Kempen, G., Olsthoorn, N., & Sprenger, S. (2011). Grammatical workspace sharing during language production and language comprehension: Evidence from grammatical multitasking. *Language and Cognitive Processes*, (October), 1–36. https://doi.org/10.1080/01690965.2010.544583

Kemper, S., Herman, R., & Lian, C. H. T. (2003). The costs of doing two things at once for young and older adults: Talking while walking, finger tapping, and ignoring speech or noise. *Psychology and Aging*, *18*(2), 181–192. https://doi.org/10.1037/0882-7974.18.2.181

Kemper, S., Schmalzried, R., Herman, R., & Mohankumar, D. (2011). The Effects of Varying Task Priorities on Language Production by Young and Older Adults. *Experimental Ageing Research*, *37*(2), 198–219. https://doi.org/10.1080/0361073X.2011.554513

Kendon, A. (1967). Some functions of gaze direction in social interaction. *Acta Psychologica, 26,* 22-63. https://doi.org/10.1016/0001-6918(67)90005-4

Kiefer, M., & Martens, U. (2010). Attentional Sensitization of Unconscious Cognition : Task Sets Modulate Subsequent Masked Semantic Priming. *Journal of Experimental Psychology. General*, *139*(3), 464–489. https://doi.org/10.1037/a0019561

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: the influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology. General*, *135*(1), 12–35. https://doi.org/10.1037/0096-3445.135.1.12

Konopka, A. E. (2011). Planning ahead: How recent experience with structures and words changes the scope of linguistic planning. *Journal of Memory and Language*, *66*(1), 143–162. https://doi.org/10.1016/j.jml.2011.08.003

Kubose, T. T., Bock, K., Dell, G. S., Garnsey, S. M., Kramer, A. F., & Mayhugh, J. (2006). The effects of speech production and speech comprehension on simulated driving performance. *Applied Cognitive Psychology*, *20*(1), 43–63. https://doi.org/10.1002/acp.1164

Kuchinke, L., Võ, M. L.-H., Hofmann, M., & Jacobs, A. M. (2007). Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *International Journal of Psychophysiology*, *65*(2), 132–140. https://doi.org/10.1016/j.ijpsycho.2007.04.004

Kuipers, J.-R., & Thierry, G. (2013). ERP-pupil size correlations reveal how bilingualism enhances cognitive flexibility. *Cortex*, Nov-Dec;*49*(10), 1–8. https://doi.org/10.1016/j.cortex.2013.01.012

Küper, K., & Heil, M. (2009). Electrophysiology reveals semantic priming at a short SOA irrespective of depth of prime processing. *Neuroscience Letters*, *453*, 107–111. https://doi.org/10.1016/j.neulet.2009.02.013

Kutas, M., Delong, K. A., & Smith, N. J. (2011). A Look around at What Lies Ahead : Prediction and Predictability in Language Processing. In M. Bar (Ed.), *Predictions in the Brain: Using our Past to Generate a Future* (pp. 190–207). Oxford: Oxford University Press. https://doi.org/10.1093/acprof

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123.Thirty

Kutas, M., & Hillyard, S. (1980). Reading Senseless Sentences: Brain Potentials Reflect Semantic Incongruity. *Science*, *207*, 203–205.

Laeng, B., Sirois, S., & Gredeback, G. (2012). Pupillometry: A Window to the Preconscious? *Perspectives on Psychological Science*, *7*(1), 18–27. https://doi.org/10.1177/1745691611427305

Laganaro, M., & Perret, C. (2011). Comparing electrophysiological correlates of word production in immediate and delayed naming through the analysis of word age of acquisition effects. *Brain Topography*, *24*(1), 19–29. https://doi.org/10.1007/s10548-010-0162-x

Lavie, N. (2005). Distracted and confused ?: Selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75-82. https://doi.org/10.1016/j.tics.2004.12.004

Lavie, N., Hirst, A., Fockert, J. W. De, & Viding, E. (2004). Load Theory of Selective Attention and Cognitive Control. *Journal of Experimental Psychology: General*, *133*(3), 339–354. https://doi.org/10.1037/0096-3445.133.3.339

Levelt, W. J. M. (1989). ACL-MIT Press series in natural-language processing. *Speaking: From Intention to Articulation*. The MIT Press.

Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences*, *3*(6), 223–232. https://doi.org/10.1016/s1364-6613(99)01319-4

Levelt, W. J. M. (2001). Spoken word production : A theory of lexical access. *PNAS*, *98* (23) 13464-13471; https://doi.org/10.1073/pnas.231459498.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–38. https://doi.org/10.1017/S0140525X99001776

Levinson, S. C. (2015). Turn-taking in Human Communication – Origins and Implications for Language Processing. *Trends in Cognitive Sciences*, *20*(1), 6–14. https://doi.org/10.1016/j.tics.2015.10.010

Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, *6*(JUN), 1–17. https://doi.org/10.3389/fpsyg.2015.00731

Lien, M., Ruthruff, E., Cornett, L., Goodin, Z., & Allen, P. A. (2008). On the Nonautomaticity of Visual Word Processing : Electrophysiological Evidence That Word Processing Requires Central Attention. *Journal of Experimental Psychology. Human Perception and Performance*, *34*(3), 751–773. https://doi.org/10.1037/0096-1523.34.3.751

Local, J., Kelly, J., Wells, B. (1986). Towards a Phonology of Conversation: Turntaking in Tyneside English. *Journal of Linguistics*, *22*(2), 411–437. Retrieved from http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=EJ347405&ERICExtSearch_SearchType_0=no&accno=EJ347405

Luck, S. J. (1998). Sources of dual-task interference: Evidence From Human Electrophysiology. *Psyhological Science*, *9*(3), 223–227. https://doi.org/10.1111/1467-9280.00043

Luck, S. J., Vogel, E. K., & Shapiro, K. L. (1996). Word meanings can be accessed but not reported during the attentional blink. *Nature*, *383*, 616–618.

MacGregor, L. J., Corley, M., & Donaldson, D. I. (2010). Listening to the sound of silence: disfluent silent pauses in speech have consequences for listeners. *Neuropsychologia*, *48*(14), 3982–3992. https://doi.org/10.1016/j.neuropsychologia.2010.09.024

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide.* (2nd ed.). Erlbaum, Mahwah, NJ.

Magyari, L., Bastiaansen, M., de Ruiter, J. P., & Levinson, S. C. (2014). Early Anticipation Lies behind the Speed of Response in Conversation. *Journal of Cognitive Neuroscience*, *26*(11), 2530–2539. https://doi.org/10.1162/jocn

Magyari, L., & de Ruiter, J. P. (2008). Timing in conversation: The anticipation of turn endings. In *12th Workshop on the Semantics and Pragmatics of Dialogue* (pp. 139–146). Retrieved from http://www.kcl.ac.uk/content/1/c6/03/89/15/papermagyari.pdf

Magyari, L., de Ruiter, J. P., & Levinson, S. C. (2017). Temporal preparation for speaking in question-answer sequences. *Frontiers in Psychology*, *8*(FEB). https://doi.org/10.3389/fpsyg.2017.00211

Maris, E. (2012). Statistical testing in electrophysiological studies. *Psychophysiology*, *49*, 549–565. https://doi.org/10.1111/j.1469-8986.2011.01320.x

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data, *164*, 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Martin, R. C., Crowther, J. E., Knight, M., Tamborello, F. P., & Yang, C.-L. (2010). Planning in sentence production: evidence for the phrase as a default planning scope. *Cognition*, *116*(2), 177–192.

https://doi.org/10.1016/j.cognition.2010.04.010

Mattys, S. L., Brooks, J., & Cooke, M. (2009). Recognizing speech under a processing load: dissociating energetic from informational factors. *Cognitive Psychology*, *59*(3), 203–243. https://doi.org/10.1016/j.cogpsych.2009.04.001

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*(7–8), 953–978. https://doi.org/10.1080/01690965.2012.705006

Mattys, S. L., & Wiget, L. (2011). Effects of cognitive load on speech recognition. *Journal of Memory and Language*, *65*(2), 145–160. https://doi.org/10.1016/j.jml.2011.04.004

Mccarthy, G., & Nobre, A. C. (1993). Modulation of semantic processing by spatial selective attention . *Electroencephalography and Clinical Neurophysiology*, *88*(3), 210–219. https://doi.org/10.1016/0168-5597(93)90005-a

Menenti, L., Gierhan, S. M., Segaert, K., & Hagoort, P. (2011). Shared Language: Overlap and Segregation of the Neuronal Infrastructure for Speaking and Listening Revealed by Functional MRI. *Psychological Science, 22*(9), 1173-82. https://doi.org/10.1177/0956797611418347

Menenti, L., Pickering, M. J., & Garrod, S. C. (2012). Toward a neural basis of interactive alignment in conversation. *Frontiers in Human Neuroscience*, *6*(June), 185. https://doi.org/10.3389/fnhum.2012.00185

Meyer, A. S., Ouellet, M., & Häcker, C. (2008). Parallel processing of objects in a naming task. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *34*(4), 982–987. https://doi.org/10.1037/0278-7393.34.4.982

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review*, *104*(1), 3–65. https://doi.org/10.1037/0033-295x.104.1.3

Meyer, P., Mecklinger, A., & Friederici, A. D. (2007). Bridging the gap between the semantic N400 and the early old / new memory effect. *NeuroReport*, *18*(10), 1009–1013. https://doi.org/10.1097/WNR.0b013e32815277eb

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*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2013). The "subjective" pupil old/new effect: is the truth plain to see? *International Journal of Psychophysiology*, *89*(1), 48–56. https://doi.org/10.1016/j.ijpsycho.2013.05.001

Naveh-benjamin, M., Craik, F. I. M., Gavrilescu, D., & Anderson, N. D. (2000). Asymmetry between encoding and retrieval processes : Evidence from divided attention and a calibration analysis. *Memory & Cognition*, *28*(6), 965–976. https://doi.org/10.3758/BF03209344

Naveh-Benjamin, M., Guez, J., & Sorek, S. (2007). The Effects of Divided Attention on Encoding Processes in Memory : Mapping the Locus of Interference. *Can J Exp Psychol.*, *61*(1), 1–12. https://doi.org/10.1037/cjep2007001.

Navon, D., & Miller, J. (2002). Queuing or Sharing? A Critical Evaluation of the

Single-Bottleneck Notion. *Cognitive Psychology*, *44*, 193–251. https://doi.org/10.1006/cogp.2001.0767

Neville, H. J., Kutas, M., Chesney, G., & Schmidt, A. L. (1986). Event-related brain potentials during initial encoding and recognition memory of congruous and incongruous words. *Journal of Memory and Language*, *25*(1), 75–92. https://doi.org/10.1016/0749-596X(86)90022-7

Olichney, J. M., Petten, C. Van, Paller, K. A., Salmon, D. P., Iragui, V. J., Kutas, M., & Olichney, C. J. M. (2000). Word repetition in amnesia. Electrophysiological measures of impaired and spared memory. *Brain*, *123*(9), 1948–1963. https://doi.org/10.1093/brain/123.9.1948

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. (2011). FieldTrip : Open Source Software for Advanced Analysis of MEG , EEG , and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*. https://doi.org/10.1155/2011/156869

Otero, S. C., Weekes, B. S., & Hutton, S. B. (2011). Pupil size changes during recognition memory. *Psychophysiology*, *48*(10), 1346–1353. https://doi.org/10.1111/j.1469-8986.2011.01217.x

Otten, L. J., Rugg, M. D., & Doyle, M. C. (1993). Modulation of event-related potentials by word repetition: The role of visual selective attention. *Psychophysiology*, *30*, 559–571. https://doi.org/10.1111/j.1469-8986.1993.tb02082.x

Papesh, M. H., & Goldinger, S. D. (2012). Pupil-BLAH-metry: cognitive effort in speech planning reflected by pupil dilation. *Attention, Perception & Psychophysics*, *74*(4), 754–765. https://doi.org/10.3758/s13414-011-0263-y

Papesh, M. H., Goldinger, S. D., & Hout, M. C. (2012). Memory strength and specificity revealed by pupillometry. *International Journal of Psychophysiology*, *83*(1), 56–64. https://doi.org/10.1016/j.ijpsycho.2011.10.002

Pashler, H. (1984). Processing Stages in Overlapping Tasks : Evidence for a Central Bottleneck. *Journal of Experimental Psychology. Human Perception and Performance*, *10*(3), 358–377. https://doi.org/10.1037//0096-1523.10.3.358

Pashler, H. (1994). Dual-Task Interference in Simple Tasks : Data and Theory. *Psychological Bulletin*, *116*(2), 220–244.

https://doi.org/10.1037/0033-2909.116.2.220.

Peelle, J. E. (2018). Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. *Ear and Hearing*, *39*(2), 204–214. https://doi.org/10.1097/AUD.0000000000000494

Phillips, N. A., & Lesperance, D. (2003). Breaking the Waves: Age Differences in Electrical Brain Activity When Reading Text With Distractors. *Psychology and Aging*, *18*(1), 126–139. https://doi.org/10.1037/0882-7974.18.1.126

Piai, V., Roelofs, A., & Maris, E. (2014). Oscillatory brain responses in spoken word production reflect lexical frequency and sentential constraint. *Neuropsychologia*, *53*, 146–156. https://doi.org/10.1016/j.neuropsychologia.2013.11.014

Pichora-fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., … Wingfield, A. (2016). Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and*

Hearing, *37*(July). https://doi.org/10.1097/AUD.0000000000000312

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *The Behavioral and Brain Sciences*, *27*(2), 169-190. https://doi.org/10.1017/S0140525X04000056

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *The Behavioral and Brain Sciences*, *36*(4), 329–347. https://doi.org/10.1017/S0140525X12001495

Piquado, T., Isaacowitz, D., & Wingfield, A. (2010). Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology*, *47*(3), 560–569. https://doi.org/10.1111/j.1469-8986.2009.00947.x

Pope, B., Blass, T., Cheek, J., & Bradford, N. H. (1971). Interviewer specificity in seminaturalistic interviews. *Journal of Consulting and Clinical Psychology, 36*(1), 152. https://doi.org/10.1037/h0030484

Redner, L. M., & Kusbit, G. W. (1991). Locus of the Moses Illusion: Imperfect encoding, retrieval, or match? *Journal of Memory and Language*, *30*(4), 385–406. https://doi.org/10.1016/0749-596X(91)90013-A

Riest, C., Jorschick, A. B., & de Ruiter, J. P. (2015). Anticipation in turn-taking : mechanisms and information sources. *Frontiers in Psychology*, *6*, 89. https://doi.org/10.3389/fpsyg.2015.00089

Roberts, S. G., Torreira, F., & Levinson, S. C. (2015). The effects of processing and sequence organization on the timing of turn taking : a corpus study. *Frontiers in Psychology,6*, 509. https://doi.org/10.3389/fpsyg.2015.00509

Roelofs, A., Ozdemir, R., & Levelt, W. J. M. (2007). Influences of spoken word planning on speech recognition. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*(5), 900–913. https://doi.org/10.1037/0278-7393.33.5.900

Roelofs, A., & Piai, V. (2011). Attention demands of spoken word planning: a review. *Frontiers in Psychology*, *2*(November), 307. https://doi.org/10.3389/fpsyg.2011.00307

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, *124*(2), 207–231. https://psycnet.apa.org/doi/10.1037/0096-3445.124.2.207

Rolke, B., Heil, M., Streb, J., & Hennighausen, E. (2001). Missed prime words within the attentional blink evoke an N400 semantic priming effect. *Psychophysiology*, *38*(2), 165–174. https://psycnet.apa.org/doi/10.1017/S0048577201991504

Rommers, J., Meyer, A. S., Praamstra, P., & Huettig, F. (2013). The contents of predictions in sentence comprehension: Activation of the shape of objects before they are referred to. *Neuropsychologia*, *51*(3), 437–447. https://doi.org/10.1016/j.neuropsychologia.2012.12.002

Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*(4), 696–735. Retrieved from http://www.jstor.org/stable/10.2307/412243

Schacter, D. L. (1987). Implicit Memory: History and Current Status. *Journal of Experimental Psychology Learning Memory and Cognition*, *13*(3), 501–518. https://psycnet.apa.org/doi/10.1037/0278-7393.13.3.501

Schaffer, D. (1983). The role of intonation as a cue to turn taking in conversation. *Journal of Phonetics*, *11*(3), 243–57. https://doi.org/10.1016/S0095-4470(19)30825-3

Schegloff, E. (1996). Turn organization: one intersection of grammar and interaction. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and Grammar* (pp. 52–133). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511620874

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, *29* (1), 86–102. https://psycnet.apa.org/doi/10.1016/0749-596X(90)90011-N

Scott, S. K., Mcgettigan, C., & Eisner, F. (2009). A little more conversation, a little less action - - candidate roles for the motor cortex in speech perception. *Nature Reviews. Neuroscience*, *10*(April), 295–302. https://doi.org/10.1038/nrn2603

Segaert, K., Menenti, L., Weber, K., Petersson, K. M., & Hagoort, P. (2012). Shared syntax in language production and language comprehension-An fMRI study. *Cerebral Cortex*, *22*(7), 1662–1670. https://doi.org/10.1093/cercor/bhr249

Severens, E., Van Lommel, S., Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica*, *119*(2), 159–187. https://doi.org/10.1016/j.actpsy.2005.01.002

Siegman, A. W., and Pope, B. (1972). The effects of ambiguity and anxiety on interviewee verbal behavior. In B. Siegman, A. W., and Pope (Ed.), *Studies in dyadic communication.* New York: Pergamon Press.

Sjerps, M. J., & Meyer, A. S. (2015). Variation in dual-task performance reveals late initiation of speech planning in turn-taking. *Cognition*, *136*, 304–324. https://doi.org/10.1016/j.cognition.2014.10.008

Skaftnes, C. (2012). *Contextual Modulation of Attentional Control Mechanisms: A Pupillometry Study (Doctoral dissertation)*. University of Oslo.

Slevc, L. R. (2011). Saying what's on your mind: working memory effects on sentence production. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *37*(6), 1503–1514. https://doi.org/10.1037/a0024350

Smith, M., & Wheeldon, L. (1999). High level processing scope in spoken sentence production. *Cognition*, *73*(3), 205–246. https://doi.org/10.1016/s0010-0277(99)00053-0

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology. Human Learning and Memory*, *6*(2), 174–215. https://doi.org/10.1037//0278-7393.6.2.174

Somberg, B. L., & Salthouse, T. A. (1982). Divided Attention Abilities in Young and Old Adults. *Journal of Experimental Psychology. Human Perception and Performance*, *8*(5), 651–663. https://doi.org/10.1037//0096-1523.8.5.651

Sommer, W., & Hohlfeld, A. (2008). Overlapping Tasks Methodology as a Tool for Investigating Language Perception. In: *Breznitz Z. (eds) Brain Research in Language. Literacy Studies, vol 1. Springer, Boston, MA*. Retrieved from http://link.springer.com/chapter/10.1007/978-0-387-74980-8_5

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., … Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(26), 10587–10592. https://doi.org/10.1073/pnas.0903616106

Strand, J. F., Brown, V. A., Merchant, M. B., Brown, H. E., & Smith, J. (2018). Measuring Listening Effort: Convergent Validity, Sensitivity, and Links With Cognitive and Personality Measures. *Journal of Speech, Language, and Hearing Research*, *61*(6), 1463–1486. https://doi.org/10.1044/2018_jslhr-h-17-0257

Strayer, D. L., & Johnston, W. A. (2001). Driven to Distraction: Dual-Task Studies of Simulated Driving and Conversing on a Cellular Telephone. *Psychological Science*, *12*(6), 462–466. https://doi.org/10.1111/1467-9280.00386

Surprenant, A. M. (1999). The Effect of Noise on Memory for Spoken Syllables. *International Journal of Psychology*, *34*(5–6), 328–333. https://doi.org/10.1080/002075999399648

Swets, B., Jacovina, M. E., Gerrig, R. J., & Gerrig, J. (2013). Effects of Conversational Pressures on Speech Planning. *Discourse Processes*, *50*(1), 23–51. https://doi.org/10.1080/0163853X.2012.727719

Tannen, D. (2007). *Talking Voices: Repetition, Dialogue and Imagery in Conversational Discourse*. Cambridge University Press.

Tombu, M., & Jolicœur, P. (2005). Testing the Predictions of the Central Capacity Sharing Model. *Journal of Experimental Psychology. Human Perception and Performance*, *31*(4), 790–802. https://doi.org/10.1037/0096-1523.31.4.790

Traxler, M. J., & Foss, D. J. (2000). Effects of sentence constraint on priming in natural language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1266–1282. https://doi.org/10.1037//0278-7393.26.5.1266

Tromp, J., Hagoort, P., & Meyer, A. S. (2016). Pupillometry reveals increased pupil size during indirect request comprehension. *The Quarterly Journal of Experimental Psychology*, *69*(6), 1093–1108. https://doi.org/10.1080/17470218.2015.1065282

Vachon, F., & Jolicoeur, P. (2011). Impaired semantic processing during task-set switching: evidence from the N400 in rapid serial visual presentation. *Psychophysiology*, *48*(1), 102–111. https://doi.org/10.1111/j.1469-8986.2010.01040.x

Vachon, F., & Jolicœur, P. (2012). On the automaticity of semantic processing during task switching. *Journal of Cognitive Neuroscience*, *24*(3), 611–626. https://doi.org/10.1162/jocn_a_00149

Van Petten, C. (2014). *Selective attention, processing load, and semantics: Insights from human electrophysiology.* (G. R. R. Mangun, Ed.). New York: Elsevier. https://doi.org/10.1016/B978-0-12-398451-7.00019-1

van Rijn, H., Dalenberg, J. R., Borst, J. P., & Sprenger, S. A. (2012). Pupil Dilation Co-Varies with Memory Strength of Individual Traces in a Delayed Response Paired-Associate Task. *PLoS ONE*, *7*(12). https://doi.org/10.1371/journal.pone.0051134

Võ, M. L.-H., Jacobs, A. M., Kuchinke, L., Hofmann, M., Conrad, M., Schacht, A., & Hutzler, F. (2008). The coupling of emotion and cognition in the eye: introducing the pupil old/new effect. *Psychophysiology*, *45*(1), 130–140. https://doi.org/10.1111/j.1469-8986.2007.00606.x

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2005). Pushing around the Locus of Selection : Evidence for the Flexible-selection Hypothesis. *Journal of Cognitive Neuroscience*, *17*(12), 1907–1922.

Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(2), 423–440. https://doi.org/10.1037/a0018619

Wang, L., Bastiaansen, M., Yang, Y., & Hagoort, P. (2011). The influence of information structure on the depth of semantic processing: how focus and pitch accent determine the size of the N400 effect. *Neuropsychologia*, *49*(5), 813–820. https://doi.org/10.1016/j.neuropsychologia.2010.12.035

Wheeldon, L. (2012). Producing spoken sentences: The scope of incremental planning. In: S. Fuchs, M. Weirich, D. Pape & P. Perrier (Eds) *Speech production and perception volume1: Speech planning and dynamics*. 97-118. Germany: Peter Lang

Wheeldon, L., & Lahiri, A. (1997). Prosodic units in speech production. *Journal of Memory and Language*, *37*(3), 356–381. https://doi.org/10.1006/jmla.1997.2517

Wickens, T. (2001). *Elementary signal detection theory*. Oxford University Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=s3pGN_se4v0C&oi=fnd&pg=PA3&dq=Elementary+Signal+Detection+Theory&ots=RYUSQzkZlY&sig=TAiNBNIg3rT-MXEDpyrRpdQOSeM

Wierda, S. M., van Rijn, H., Taatgen, N. A., & Martens, S. (2012). Pupil dilation deconvolution reveals the dynamics of attention at high temporal resolution. *Proceedings of the National Academy of Sciences*, *109*(22), 8456–8460. https://doi.org/10.1073/pnas.1201858109

Wilson, M., & Wilson, T. P. (2005). An oscillator model of the timing of turn-taking. *Psychonomic Bulletin & Review*, *12*(6), 957–968. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16615316

Winn, M. B., & Moore, A. N. (2018). Pupillometry Reveals That Context Benefit in Speech Perception Can Be Disrupted by Later-Occurring Sounds, Especially in Listeners With Cochlear Implants. *Trends in Hearing*, *22*, 1–22. https://doi.org/10.1177/2331216518808962

Wlotko, E. W., Federmeier, K. D., & Kutas, M. (2012). To predict or not to predict: Age-related differences in the use of sentential context. *Psychology and Aging*, *27*(4), 975–988. https://doi.org/10.1037/a0029206.To

Zekveld, A. A., Heslenfeld, D. J., Johnsrude, I. S., Versfeld, N. J., & Kramer, S. E. (2014). The eye as a window to the listening brain : Neural correlates of pupil size as a measure of cognitive listening load. *NeuroImage*, *101*, 76–86. https://doi.org/10.1016/j.neuroimage.2014.06.069

APPENDIX

A. Picture Names of chapter 2 with English translations in parenthesis

aap (monkey), ananas (pineapple), appel (apple), auto (car), badkuip (bathtub), ballon (balloon), bank (couch), barbeque (barbeque), bed (bed), beker (trophy), bel (bell), blik (can), bloem (flower), blouse (Blouse), boekenkast (bookcase), boerderij (farm), boom (tree), boot (boat), brandweerman (fireman), brievenbus (mailbox), broek (pants), brood (bread), broodrooster (toaster), brug (bridge), bureau (bureau), bus (bus), cactus (cactus), camera (camera), caravan (caravan), clown (clown), cocktail (cocktail), computer (computer), dak (roof), das (tie), dokter (doctor), doos (box), draak (dragon), drumstel (drums), eekhoorn (squirrel), eend (duck), eenhoorn (unicorn), eiland (island), emmer (bucket), fabriek (factory), fiets (bicycle), fornuis (stove), geit (goat), gieter (wateringcan), giraffe (giraffe), gitaar (guitar), glas (glass), glijbaan (slide), grasmaaier (lawnmower), hagedis (lizard), handtas (purse), hart (heart), heks (witch), helm (helmet), hemd (shirt), hond (dog), huis (house), hut (hut), iglo (igloo), indiaan (indian), jas (jacket), jeep (jeep), jongen (boy), jurk (dress), kaars (candle), kado (present), kan (pitcher), kangoeroe (kangaroo), kano (canoe), kanon (cannon), kast (dresser), kasteel (castle), kat (cat), katapult (slingshot), kerk (church), kikker (frog), kip (chicken), kleerkast (closet), koe (cow), koets (coach), konijn (rabbit), koning (king), kopje (cup), kraan (faucet), krokodil (alligator), kroon (crown), kruiwagen (wheelbarrow), kruk (stool), laars (boot), lamp (lamp), leeuw (lion), man (man), mand (basket), masker (mask), meisje (girl), microscoop (microscope), mixer (mixer), molen (windmill), neushoorn (rhinoceros), ober (waiter), octopus (octopus), olifant (elephant), paard (horse), paddestoel (mushroom), paleis (palace), palmboom (palmtree), pan (pot), pantoffel (slipper), papegaai (parrot), parachute (parachute), pauw (peacock), pet (hat), piano (piano), piraat (pirate), plant

(plant), pomp (bicyclepump), pop (doll), priester (priest), revolver (gun), ridder (knight), robot (robot), rok (skirt), rolschaa (rollerskate), rolstoel (wheelchair), rugzak (backpack), schaap (sheep), schaar (scissors), schilderij (picture), schildpad (turtle), schoen (Shoe), schommel (swing), schommelstoel (rockingchair), skateboard (skateboard), slak (snail), slee (sled), sok (sock), soldaat (soldier), spaghetti (spaghetti), spiegel (mirror), spin (Spider), spook (ghost), step (step), stoel (chair), stofzuiger (vacuum), stoplicht (Traffic light), strijkijzer (Iron), taart (cake), tafel (table), tank (tank), telescope (telescope), tent (tent), tevee (Television), tijger (tiger), tractor (tractor), trap (stairs), trein (train), trommel (drum), trui (sweater), vaas (vase), varken (pig), vergiet (drainer), verpleegster (nurse), vliegtuig (airplane), voet (foot), vos (fox), vrachtwagen (truck), vrouw (woman), vuur (fire), weegschaal (scale), wekker (alarmclock), wieg (crib), zadel (saddle), zak (sack), zebra (zebra), zoutvat (salt), zwaan (swan)

B. Auditory target words of chapter 3 in Dutch with English translations in parenthesis.

baat (benefit), bad (bath), bal (ball), bank (bench/bank), blad (leaf), blok (block), bocht (curve), boord (collar), bos (woods), bril (glasses), broer (brother), bron (source), brood (bread), brug (bridge), bui (shower), buurt (neighbourhood), cel (cell), club (club), daad (deed), dal (valley), deur (door), dief (thief), doos (box), droom (dream), eind (island), fles (bottle), gang (hallway), gat (hole), geld (money), glas (glass), haar (hair), hals (neck), hart (hart), heer (gentleman), hek (fence), hoek (corner), hond (dog), hout (wood), hut (hut), jas (coat), jurk (dress), kaas (cheese), kans (chance), keel (throat), kier (crack), kist (chest), klas (classroom), knie (knee), koor (choir), lamp (lamp), last (burden), lied (song), lift (elevator), lot (fate), maan (moon), macht (power), mes (knife), moed (courage), mond (mouth), neef (cousin), paard (horse), pan (pan), pech (bad luck), pil (pill), plek (place), prijs (price), puin (debris), punt (point), raam (window), reeks (series), rest (rest), reus (giant), riet (reed), roem (fame), rust (rest), schok (shock), sneeuw (snow), som (sum), stad (city), stel (couple), stoel (chair), stok (stick), taal (language), tent (tent), top (top), touw (rope), tuin (garden), vel (skin), vloer (floor), vuist (fist), wang (cheek), week (week), wieg (cradle), wijn (wine), wind (wind), wit (white), zak (bag), zand (sand), zoon (son), zout (salt).

C. Picture Names[10 (see footnotes of chapter 3)] of chapter 3 with English translations in parenthesis.

aambeeld (anvil), aardbei (strawberry), accordeon (accordion), ajuin (onion), ananas (pineapple), anker (anchor), artisjok (artichoke), ballon (balloon), banaan (banana), barbecue (barbecue), bever (beaver), bezem (broom), bokaal (jar), boormachine (drill), borstel (brush), brievenbus (mailbox), broodrooster (toaster), cactus (cactus), denneappel (pinecone), deurknop (doorknob), dinosaurus (dinosaur), dolfijn (dolphin), eikel (acorn), enveloppe (envelope), fornuis (stove), fototoestel (camera), gewei (antlers), gieter (wateringcan), gitaar (guitar), glijbaan (slide), hagedis (lizard), helicopter (helicopter), hengel (fishingpole), hoefijzer (horseshoe), iglo (igloo), ijsje (icecreamcone), jojo (yoyo), kangoeroe (kangaroo), kapstok (hanger), katapult (slingshot), ketting (necklace), kever (bug), kikker (frog), kinderwagen (stroller), kleerkast (closet), koning (king), kruiwagen (wheelbarrow), kurketrekker (corkscrew), lippen (lips), magneet (magnet), mixer (mixer), olifant (elephant), paddestoel (mushroom), palmboom (palmtree), paperclip (clip), parachute (parachute), paraplu (umbrella), pelikaan (pelican), penseel (paintbrush), piramide (pyramid), platenspeler (recordplayer), pleister (bandaid), pompoen (pumpkin), puzzel (puzzle), racket (rocket), robot (robot), rolschaats (rollerskate), scheermes (razor), schommel (swing), selder (celery), slijper (pencilsharpener), sneeuwman (snowman), spaarvarken (piggybank), spinneweb (spiderweb), statief (tripod), strijkijzer (iron), struisvogel (ostrich), tandenborstel (toothbrush), tennisracket (tennisracket), tomaat (tomato), tractor (tractor), trechter (funnel), trommel (drum), trompet (trumpet), tuinslang (hose), ventilator (fan), verkeerslicht (stoplight), verrekijker (telescope), vingerhoed (thimble), vlieger (kite), vliegtuig (airplane), vlinder (butterfly), vrachtwagen (truck), wandelstok (cane), wasknijper (clothespin), weegschaal (scale), wereldbol (globe), zaklamp (flashlight), zeehond (seal), zeilboot (sailboat).

D. Lure words of chapter 3 in Dutch, with English translations in parenthesis.

Arts (doctor), baan (job), bass (boss), bar (bar), beek (brook), beeld (picture), beest (beast), berg (mountain, bier (beer), boog (bow), boom (tree), boot (boat), brief (letter), broek (trousers), buik (belly), bus (bus), dak (roof), dier (animal), ding (thing), doek (cloth), doel (goal), dorp (village), eed (oath), fiets (bicycle), film (movie), flat (flat), fout (error), gas (gas), geest (spirit), geur (scent), goud (gold), gras (grass), groep (group), hal (hall), ham (ham), held (hero), hemd (vest), hoed (hat), hoorn (horn), jacht (hunting), kaart (card), kas (cash-register), kast (closet), kerk (church), kern (core), keus (choice), kin (chin), klok (clock), koers (course), kring (circle), laan (avenue), les (lesson), licht (light), lijf (body), lijn (line), lijst (list), lucht (air), maand (month), muur (wall), nacht (night), neus (nose), nood (need), park (park), pijn (pain), plaat (plate), plein (square), poort (gate), pot (jar), rand (edge), rijk (rich), ring (ring), rok (skirt), rug (back), sap (juice), schip (ship), school (school), sfeer (atmosphere), soep (soup), spel (game), stam (stem), steen (stone), tas (purse), term (terminology), ton (barrel), toon (tone), trein (train), tuig (rig), vak (course), veld (field), vis (fish), vlees (meat), voet (foot), volk (people), vuur (fire), wiel (wheel), wol (wool), woud (forest), zeep (soap), zicht (sight), ziel (soul).

E. The 200 constraining and not constraining Dutch sentence-frames of chapter 4 (with approximate English translations underneath) in which the 100 sentence-final-words were embedded, resulting in predictable (P) and unpredictable (UP) versions of the sentence-final-words.

Nr.	Sentence frame with sentence-final-word	condition
1	De Nachtwacht van Rembrandt is een bekend schilderij The 'Night Watch' by Rembrandt is a famous painting.	P
	De reclame ging over een duur schilderij The advertisement was about an expensive painting	UP
2	Haar moeder is ziek maar wil niet naar de dokter Her mother is ill, but (she) does not want to go to the doctor	P
	De man had slechts één dokter The man had only one doctor	UP
3	Omdat ze zulke koude oren had droeg het meisje een muts Because she had such cold ears, the girl wore a hat	P
	De posters langs de weg maken veel reclame voor een muts The posters along the road advertised a lot for a hat	UP
4	Ik moet nieuwe snaren kopen voor mijn gitaar I have to buy new strings for my guitar	P
	Het werd een oude gitaar It was getting an old guitar	UP
5	Ze kunnen de rivier niet over vanwege een reparatie aan de brug They cannot cross the river because of a reparations to the bridge	P
	Ze kunnen hier niet lopen vanwege een onbekend probleem aan de brug They cannot walk here because of an unknown problem with the bridge	UP
6	Deze twee planken moet je vast zetten met een schroef You have to fasten these two planks with a screw	P
	Dit probleem moet je kunnen oplossen met een schroef You have to be able to fix this problem with a screw	UP
7	Het regende maar gelukkig had ik een paraplu It was raining, but fortunately I had an umbrella	P
	Je vergeet het maar waarschijnlijk heb ik een paraplu You are forgetting it, but I probably have an umbrella	UP
8	Bugs Bunny kauwde lekker op een wortel Bugs Bunny was chewing nicely on a carrot	P
	We eten graag een wortel We like to eat a carrot	UP
9	Om de stroom uit te zetten, drukte hij op de schakelaar To turn the power off, he pushed the switch	P
	Om alles in orde te kunnen krijgen, gebruikte hij een schakelaar To make everything (all) right, he used a switch	UP
10	Voor zijn kano gebruikt de indiaan een peddel For his canoe, the Indian uses a paddle	P

Nr.	Sentence frame with sentence-final-word	condition
	Op straat horen ze toevallig over een peddel On the streets, they coincidentally hear about a paddle	UP
11	Om zijn nek te beschermen, droeg hij een sjaal To protect the back of his neck, he wore a scarf	P
	Het meisje maakte een sjaal The girl made a scarf	UP
12	Om zijn kale hoofd te beschermen, draagt hij een mooie pet To protect his bald head, he wears a nice cap	P
	De winkel op de hoek verkoopt een mooie pet The store on the corner sells a nice cap	UP
13	We staken het vuur aan met een lucifer We started the fire using a match	P
	We vroegen de mensen om een lucifer We asked people for a match	UP
14	De Utrechtse Dom luidt heel hard haar klok The Cathedral of Utrecht rings very loudly her bells	P
	Ik word helemaal gek van het geluid van de klok I go mad from the sound of the bells	UP
15	Het schip werd stevig vastgelegd met een anker The ship was firmly established with an anchor	P
	Dat zware metalen stuk daar is een anker This heavy metal thing there is an anchor	UP
16	Mijn favoriete zomerdessert is natuurlijk een ijsje My favorite summerdessert is of course ice cream	P
	Ze liepen gisteren voorbij met een ijsje They walked by yesterday with ice cream	UP
17	Je moet altijd kloppen op mijn deur You should always knock on my door	P
	Zij verbaast zich nog steeds over de deur She is still amazed by the door	UP
18	Het vogeltje kon niet opstijgen vanwege zijn gebroken vleugel The bird could not fly away because of his broken wing	P
	Het object had de vorm van een vleugel The object had the shape of a wing	UP
19	Om een rechte lijn te trekken, gebruik ik een liniaal To draw a straight line, I use a ruler	P
	Om dit goed te kunnen, gebruik ik een liniaal To do this right, I use a ruler	UP
20	In de wind wapperde buiten een vlag Outside in the wind blew a flag	P
	In deze stad zie je overal buiten een vlag In this town you see everywhere oustide a flag	UP
21	In die fabriek bottelt men wijn in verschillende soorten flessen In that factory, people bottle wine in different types of bottles	P
	De vader vraagt de jonge man naar verschillende soorten flessen The fathers asks the young man about different types of bottles	UP
22	Ik houd van kerst want ik krijg altijd een kado I love christmas because I always get a present	P
	Morgen koop ik voor haar alleen een kado Tomorrow, I will buy only for her a present	UP
23	Vanwege die flat wast de glazenwasser iedere maand de ramen Because of that apartment, the window-cleaner washed every month the windows	P
	Wat de vreemde man probeert te beschrijven zijn de ramen What the weird man tries to describe, are the windows	UP
24	Op mijn broodje smeer ik altijd eerst een laag boter On my bread, I always spread first a layer of butter	P

Nr.	Sentence frame with sentence-final-word	condition
	Na lang nadenken kiest ze toch voor een beetje boter After giving it much thought, she chooses (after all) for a little butter	UP
25	De kok sneed de courgette op een snijplank The cook cut the zucchini on a chopping board	P
	De vrouw is op zoek naar een snijplank The woman is looking for a chopping bord	UP
26	De boer melkte de koe The farmer milked the cow	P
	Het kind tekende een koe The child drew a cow	UP
27	Voor het kinderfeest hebben we al een opgeblazen ballon For the children's party we already have an inflated balloon	P
	Voor de show hebben we al een goede ballon For the show we already have a good balloon	UP
28	Ik wil drinken vanavond dus ik ga naar de kroeg met de fiets I want to drink tonight, so I go to the pub by bike	P
	Ik kwam gisteren thuis en toen zag ik buiten nog een fiets I came home yesterday and I saw outside a(nother) bike	UP
29	Hij sloeg op zijn duim met een hamer He hit his thumb with a hammer	P
	Uit de grote lade haalde hij een hamer Out of the big drawer, he got a hammer	UP
30	In het park voerden de kinderen de eend In the park, the children fed the duck	P
	Zij werkt aan een verhaal over de eend She is working on a story about the duck	UP
31	Ze bedekte de snee in haar vinger met een pleister She covered the cut in her finger with a band-aid	P
	Ze doorzocht het vak in de winkel voor een pleister She searched the shelves in the shop for a band-aid	UP
32	Ik schep meestal teveel op mijn bord I usually dish up too much (on my plate)	P
	Ik doe vaak veel op het bord I often put a lot on the plate	UP
33	Het slijmspoor op de aangevreten sla in de moestuin kwam door een slak The slimetrace on the gnawed at lettuce in the vegetable garden was caused by a snail	P
	In die zaal wordt er vandaag een korte voorlichting gegeven over een slak In that hall, short information will be given today about a snail	UP
34	Vanwege ontstekingen trok de tandarts haar kies Because of inflammations, the dentist extracted her molar	P
	Normaal gesproken gebeurt dat nooit met zijn kies Normally, that never happens to his molar	UP
35	Hij veegde de vloer met een bezem He sweeped the floor with a broom	P
	Hij zag de tekening van een bezem He saw the drawing of a broom	UP
36	Spanje is een land met veel zon Spain is a country with lots of sunshine	P
	Dit is een plaats met veel zon This is a place with lots of sunshine	UP
37	Voor de verlichting naast mijn bed heb ik een lamp For the lighting next to my bed, I have a lamp	P
	Voor het bezoek van mijn oom kocht ik een lamp For the visit to my uncle, I bought a lamp	UP

Nr.	Sentence frame with sentence-final-word	condition
38	De chauffeur stempelde mijn strippenkaart in de bus The driver stamped my card in the bus	P
	Vandaag schrijft hij een werkstuk over de bus Today, he writes an essay about the bus	UP
39	Pinoccio was een jongen met een lange neus Pinocchio was a boy with a long nose	P
	Hein was een jongen met een lange neus Hein was a boy with a long nose	UP
40	De stroom viel uit maar gelukkig had ik thuis nog een kaars The power went out, but luckily I still had at home a candle	P
	Zij was wel blij want gelukkig had ik thuis nog een kaars She was happy, because luckily I had at home a candle	UP
41	Op zondagen hang ik vaak lekker op de bank On sundays, I often hang around nicely on the couch	P
	Tante Marga was heel blij met haar bank Aunt Marga was very happy with her couch	UP
42	Aan de logeerders geef ik liever een zacht kussen To the visitors I'd rather give a soft pillow	P
	Aan de jongens geef ik liever een slecht kussen To the boys I'd rather give a bad pillow	UP
43	Het verkeer bij de kruising stond te wachten voor een stoplicht The traffic at the intersection was waiting for a traffic light	P
	De vrouw wil heel graag iets kunnen vertellen over een stoplicht The woman really wants to be able to tell something about a traffic light	UP
44	Ze bewaart haar lipstift en maskara in haar tas She keeps her lipstick and her mascara in her purse	P
	Het dure tijdschrift adverteerde met een dure tas The expensive magazine advertised an expensive purse	UP
45	Vroeger moesten mensen water halen uit een put Back in the days, people had to get water out of a well	P
	Vroeger moesten mensen lang lopen naar een put Back in the days, people had to walk far for a well	UP
46	Door het zand sneed de slipper tussen zijn tenen Because of the sand, the flip-flop cut between his toes	P
	Het meisje merkte niets aan haar tenen The girl did not notice anything about her toes	UP
47	We rijden graag dus we kochten pas een auto We like to drive, so we recently bought a car	P
	Op een gewone dag gebruik ik liever geen auto On a normal day, I'd rather not use a car	UP
48	Zij scheidde netjes haar haar met een kam She parted her hair neatly with a comb	P
	Voor de jonge vrouw pakken we de kam For the young woman, we take a comb	UP
49	Soep eet je met een lepel You eat soup with a spoon	P
	We proberen het eerst met een lepel First, we try it with a spoon	UP
50	Carla bewaart haar oude poppen in een houten kist Carla keeps her old dolls in a wooden box	P
	Op dit moment bevinden ze zich in een kist At the moment they, are in a box	UP
51	Jans ogen gingen tranen door het snijden van een ui Jan's eyes started watering by cutting an onion	P
	Heel voorzichtig raakt het meisje met haar vinger een ui Very carefully, the girl touched with her finger an onion	UP
52	De Egyptenaren bouwden een grote pyramide The Egyptians built a large pyramid	P

Nr.	Sentence frame with sentence-final-word	condition
	De ouderen bouwden een grote pyramide The elderly people built a large pyramid	UP
53	De vijand van Peter Pan heeft geen hand maar een haak Peter Pan's enemy does not have a hand, but a hook.	P
	Ze wilden iets anders dus ze kochten een haak They wanted something different, so they bought a hook	UP
54	Haar walkman deed het niet meer vanwege de batterij Her walkman did not work anymore, because of the battery	P
	In die plastic zak zit nog een batterij In that plastic bag is still a battery	UP
55	Het huis van een eskimo is een iglo An Eskimo's house is an iglo	P
	Het huis van Anna Chena is een iglo Anna Chena's house is an iglo	UP
56	Ik stop de brief in een gekleurde envelope I put the letter in the colored envelope	P
	Ik doe dit hier in de correcte envelope I put this (here) in the right envelope	UP
57	Tijdens de kampeervakantie zat hij nauwelijks in zijn tent During the campingholiday, he was barely in his tent	P
	Gedurende de dag was hij nauwelijks in de tent During the day, he was barely in the tent	UP
58	De tuin wordt omgrensd door een hekje The garden is bordered by a small fence	P
	Toen ik daar keek, zag ik een hekje When I looked there, I saw a small fence	UP
59	Ik moet dit papier knippen maar heb helaas geen schaar I have to cut this paper, but unfortunately I don't have any scissors.	P
	Ik moet dit kunnen aanpassen maar heb helaas geen schaar I have to be able to adjust this, but unfortunately I don't have any scissors.	UP
60	De bezoekers keken naar de uitbarsting van de vulkaan The visitors watched the eruption of the volcano	P
	Mijn opa kent goede verhalen over een vulkaan My grandfather knows good stories about a volcano	UP
61	Hij kon het slot niet openen zonder de sleutel He could not open the lock without the key	P
	Hij had moeite met het tekenen van een sleutel He had trouble drawing a key	UP
62	In Gouda gebruiken ze goede melk voor hun kaas In Gouda, they use good milk for their cheese	P
	De koopman pakt snel de doos met de kaas The merchant quickly grasps the box with the cheese	UP
63	Mijn konijn zit nooit opgesloten in de kooi My rabbit is never locked up in the cage	P
	Ik las een verhaal over een kooi I read a story about a cage	UP
64	Uit de naaidoos pakte mijn oma een scherpe naald Out of the sewing-box my grandmother took a sharp needle	P
	De groep daar heeft het over een algemene naald The group over there is talking about a common needle	UP
65	Hij hakte de boom om met een bijl He cut down the tree with an axe	P
	In de reclame zag je een bijl In the advertisement, you saw an axe	UP
66	De trein rijdt doorgaans over een spoor The train usually runs on the tracks	P

Nr.	Sentence frame with sentence-final-word	condition
	Op die plek ligt soms een spoor On that spot lies sometimes a track	UP
67	De astronauten landden op de maan The astronauts landed on the moon	P
	De kinderen wezen naar de maan The children pointed at the moon	UP
68	In de keukenhof fotografeerden de japanse toeristen de tulpen In the keukenhof garden the Japanese tourists photographed the tulips.	P
	Die persoon keek nog eens goed naar de tulpen That person took another good look at the tulips	UP
69	Om plantjes water te geven gebruik je een gieter To water plants, you use a watering can	P
	Om dit voor elkaar te krijgen, gebruik je een gieter To accomplish this, you use a watering can	UP
70	De boogschieter schoot met een puntige pijl The archer shot with a sharp arrow	P
	De kerel sprong naast de bronzen pijl The big fellow jumped off the bronze arrow	UP
71	Om vliegen te vangen maakt die spin een web To catch flies, that spider makes a web	P
	Heel voorzichtig betast hij met zijn vingers een web Very carefully, he feels with his fingers a web	UP
72	Moeder maakt jus d'orange uit drie sinaasappels Mother makes orange juice out of three oranges	P
	In de kelder thuis liggen veel sinaasappels In the cellar at home there are a lot of oranges	UP
73	De valentijnskaart had de vorm van een hart The Valentine's card had the shape of a heart	P
	Het huis had de vorm van een hart The house had the shape of a heart	UP
74	De samoerai vocht met een zwaard The samurai fought with a sword	P
	De dissident was tevreden met zijn zwaard The dissident was satisfied with his sword	UP
75	Het team speelde goed en won natuurlijk de beker The team played well and won of course the cup	P
	Op de stoep vonden de jonge meisjes een beker On the sidewalk, the young girls found a cup	UP
76	Wijn wordt gemaakt van druiven Wine is made out of grapes	P
	Dit wordt gemaakt van druiven This is made out of grapes	UP
77	Kermit is een heel sympathieke kikker Kermit is a very sympathetic frog	P
	Hij luisterde even naar het geluid van de kikker He listened to the sound of the frog	UP
78	Het Carnaval van Venetie staat bekend om zijn maskers The Carnival in Venice is famous for its masks	P
	De kunstenaar van dat theater is bekend voor zijn maskers The artist from that theater is famous for his masks	UP
79	Als we naar de bioscoop gaan, eten we altijd popcorn When we go to the movie theater, we always eat popcorn	P
	Als we naar huis gaan, maken we altijd popcorn When we go home, we always make popcorn	UP
80	Om de kamer donker te maken, heb ik zwarte gordijnen To make the room dark, I have black curtains	P
	De jongen kocht op de markt gele gordijnen The boy bought on the market yellow curtains	UP

Nr.	Sentence frame with sentence-final-word	condition
81	Met haar nagel trok moeder een ladder in haar panty With her nail, mother pulled a run in her tights	P
	Die mevrouw in de vorige winkel noemde dit een panty The woman in the previous shop called these tights	UP
82	Om haar vinger droeg ze een mooie ring Around her finger, she wore a beautiful ring	P
	De man kocht een mooie ring The man bought a beautiful ring	UP
83	Ze kochten een wieg voor de baby They bought a cradle for the baby	P
	Ze kochten fruit voor de baby The bought fruit for the baby	UP
84	Ze wisten niet hoe laat het was want ze hadden geen horloge They did not know what time it was, because they did not have a watch	P
	Ze waren in gevaar want ze liepen rond met een opvallend horloge They were in danger, because they walked around with a noticeable watch	UP
85	De filosoof genoot van het roken van zijn pijp The philosopher enjoyed smoking his pipe	P
	De man genoot van het ruiken aan de pijp The man enjoyed smelling the pipe	UP
86	Zijn broek zakt af want hij heeft geen riem His pants come down, becaus he does not have a belt	P
	Hij kocht in de winkel een lange riem In the store, he bought a long belt	UP
87	In de slaapkamer zwierf een irritante mug In the bedroom flew around an annoying mosquito	P
	In de ruimte was een irritante mug In the room there was an annoying mosquito	UP
88	Voor Halloween knutselden de kinderen een grote pompoen For Halloween, the children tinkered a big pumpkin	P
	Voor de maaltijd bereidde haar moeder een grote pompoen For the meal, her mother cooked a big pumpkin	UP
89	Voor de ontspanning neem ik graag een warm bad To relax, I like to take a warm bath	P
	In de avond neem ik heel graag een bad In the evenings I really like to take a bath	UP
90	Een vogel die alles herhaalt is een papegaai A bird that repeats everything is a parrot	P
	Ik hou van het geluid van een papegaai I love the sound of a parrot	UP
91	De agressieve jongen wilde vechten en balde zijn vuist The aggressive boy wanted to fight and clenched his fist	P
	Op de tentoonstelling ziet hij foto's van een vuist At the exhibition, he showed pictures of a fist	UP
92	In de wei plukte het meisje een bloem In the meadow, the girl picked a flower	P
	Ze keken rond en vonden toch een bloem They looked around and after all found a flower	UP
93	De schoonmaakster gooit water uit een emmer The cleaner throws water out of a bucket	P
	De vrouw doet dit in de emmer The woman puts this in the bucket	UP
94	Het is ongezond om alleen groente te eten uit een blik It is unhealthy to only eat vegetables out of a can	P
	Het is niet goed om iets te halen uit een blik It is not right to get something out of a can	UP

Nr.	Sentence frame with sentence-final-word	condition
95	Ik hang heel graag voor de tv I really like to hang in front of the tv	P
	In de nieuwe reclame zag je een tv In the new advertisement, you saw a tv	UP
96	Mijn opa kan niets lezen zonder zijn bril My grandfather cannot read anything without his glasses	P
	Ik moet spoedig zorgen voor een nieuwe bril I have to get myself soon some new glasses	UP
97	De wijnfles zat nog dicht met een kurk The winebottle was still closed with a cork	P
	Ik maak dit nog dicht met een kurk I will close this with a cork	UP
98	In de lente zet ik bloembakken op het balkon In spring, I put flower boxes on the balcony	P
	Ik heb mijn broer gevraagd mee te helpen aan het balkon I asked my brother to help with the balcony	UP
99	Om de radio uit te zetten, trok hij aan de stekker To turn the radio off, he pulled the plug	P
	Hij liep naar de winkel en kocht een stekker He walked to the store and bought a plug	UP
100	Voor Sinterklaas zetten de kinderen hun schoen For 'Saint Nicholas/Sinterklaas' children put their shoe (next to the chimney)	P
	Op de grond zetten de kinderen hun schoen The children put on the ground their shoe	UP

F. Auditory target words (sentence-final-words) of chapter 4 with English translations in parenthesis.

anker (anchor), auto (car), baby (baby), bad (bath), balkon (balcony), ballon (balloon), bank (couch), batterij (battery), beker (cup), bezem (broom), bijl (axe), blik (can), bloem (flower), bord (plate), boter (butter), bril (glasses), brug (bridge), bus (bus), deur (door), dokter (doctor), druiven (grapes), eend (duck), emmer (bucket, envelop (envelope), fiets (bike), flessen (bottles), gieter (watering can), gitaar (guitar), gordijnen (curtains), haak (hook), hamer (hammer), hart (heart), hekje (small fence), horloge (watch), iglo (iglo), ijsje (ice cream), kaars (candle), kaas (cheese), kado (present), kam (comb), kies (molar), kikker (frog), kist (box), klok (clock), koe (cow), kooi (cage), kurk (cork), kussen (pillow), lamp (lamp), lepel (spoon), liniaal (ruler), lucifer (match), maan (moon), maskers (masks), mug (mosquito), muts (hat), naald (needle), neus (nose), panty (tights), papegaai (parrot), paraplu (umbrella), peddel (paddle), pet (cap), pijl (arrow), pijp (pipe), pleister (band-aid), pompoen (pumpkin), popcorn (popcorn), put (well), pyramide (pyramid), ramen (windows), riem (belt), ring (ring), schaar (scissors), schakelaar (switch), schilderij (painting), schoen (shoe), schroef (screw), sinaasappels (oranges), sjaal (scarf), slak (snail), sleutel (key), snijplank (chopping board), spoor (track), stekker (plug), stoplicht (traffic light), tas (purse), tenen (toes), tent (tent), tulpen (tulips), tv (tv), ui (onion), vlag (flag), vleugel (wing), vuist (fist), vulkaan (volcano), web (web), wortel (carrot), zon (sun), zwaard (sword).

G. Picture Names of chapter 4 with English translations in parenthesis

aardbei (strawberry), accordeon (accordion), ananas (pineapple), banaan (banana), barbecue (barbecue), borstel (brush), brievenbus (mailbox), cactus (cactus), denneappel (pinecone), dinosaurus (dinosaur), dolfijn (dolphin), fornuis (stove), fototoestel (camera), gewei (antlers), glijbaan (slide), hagedis (lizard), helicopter (helicopter), hengel (fishingpole), jojo (yoyo), kangoeroe (kangaroo), katapult (slingshot), ketting (necklace), kever (bug), mixer (mixer), olifant (elephant), paddestoel (mushroom), paperclip (clip), pelikaan (pelican), puzzel (puzzle), racket (rocket), robot (robot), rolschaats (rollerskate), slijper (pencilsharpener), spaarvarken (piggybank), strijkijzer (iron), tandenborstel (toothbrush), tomaat (tomato), trechter (funnel), tuinslang (hose), ventilator (fan), verrekijker (telescope), vlieger (kite), vliegtuig (airplane), vlinder (butterfly), vrachtwagen (truck), wandelstok (cane), wasknijper (clothespin), weegschaal (scale1), wereldbol (globe), zaklamp (flashlight)

H. Lures of the recognition memory test in chapter 4 with English translations in parenthesis

arm (arm), baard (beard), bel (bell), berg (mountain), bladen (sheets/magazines/leaves), boek (book), boerderij (farm), bom (bomb), brood (bread), bruid (bride), bureau (desk), cowboy (cowboy), dak (roof), diamant (diamond), douche (shower), draak (dragon), ei (egg), eland (moose/elk), fonteinen (fountains), glas (glass), graan (grain), hals (neck), handdoek (towel), harp (harp), helmen (helmets), hemd (vest/undershirt), hoef (hoof), hond (dog), hooi (hay), houweel (pickaxe), jurk (dress), kanon (cannon), kasteel (castle), kerken (churches), kers (cherry), knoop (button), koffer (suitcase), kom (bowl), kroon (crown), laars (boot), landkaart (map), lat (slat/strip), mand (basket), map (folder), mier (ant), molen (mill), munt (coin), net (net), neushoorn (rhino), ober (waiter), oorbel (earring), orgel ((pipe) organ), orkest (orchestra), pallet (pallet(board)), pauw (peacock), pilaar (pillar), plank (shelf/plank/board), priester (priest), pruik (wig), rits (zipper), rok (skirt), rugzak (rucksack), schaap (sheep), scharnier (hinge), schoorsteen (chimney), schop (shovel), slager (butcher), soldaat (soldier), spaghetti (spaghetti), spatel (spatula), spiegel (mirror), sprinkhaan (grasshopper), ster (star), stoelen (chairs), stofzuiger (vacuum cleaner), taart (cake), tafel (table), tank (tank), tapijten (carpets), telefoon (telephone), thermometer (thermometer), touw (rope), vazen (vases), verpleegster (nurse), viool (violin), vlot (raft), vos (fox), vuurtoren (ligthouse), wasbak (sink), wasmachine (washing machine), waterput (well), wiel (wheel), wolken (clouds), worst (sausage), zaag (saw), zadel (saddle), zakdoek (handkerchief), zebra (zebra), zetel (seat), zwaan (swan).

I. Mean duration (M) in milliseconds, standard error (S.E.), t- and p-values for all the relevant comparisons within and across lists of the exposure phase in chapter 5.

	Sentence duration within				Sentence final word duration within			
	List1		List2		List1		List2	
	expected	anomalous	expected	anomalous	expected	anomalous	expected	anomalous
M (ms)	4178	4008	3974	4214	567	553	550	589
S.E.	146	145	143	143.74	16.65	13.63	13.95	17
t	.826		-1.185		.681		-1.738	
p	.41		.237		.496		.084	
	Sentence duration across				Sentence final word duration across			
	List1		List2		List1		List2	
M (ms)	4093		4094		560		569	
S.E.	103		101		10.74		11	
t	-.0081				-.608			
p	.993				.543			

J. Auditory sentences of chapter 5 in Dutch with approximate English translations in parenthesis.

Nr.	sentence	Condition
1	"Het zoete stuk fruit dat Willem Tell met pijl en boog van zijn zoontje's hoofd schoot was een appel." The sweet piece of fruit that William Tell shot from his son's head with a bow and arrow was an apple	expected
	"Het zoete stuk fruit dat Willem Tell met pijl en boog van zijn zoontje's hoofd schoot was een molen." The sweet piece of fruit that William Tell shot from his son's head with a bow and arrow was a windmill	anomalous
2	"Door de mouw van een jas steek je normaliter je arm." Through the sleeve of a coat you normally put your arm.	expected
	"Door de mouw van een jas steek je normaliter je hoop." Through the sleeve of a coat you normally put your hope.	anomalous
3	"De rijke vrouw droeg om haar pols een sierlijke armband." The rich woman was wearing around her wrist an elegant bracelet.	expected

Nr.	sentence	Condition
	"De rijke vrouw droeg om haar pols een sierlijke blocnote." The rich woman was wearing around her wrist an elegant bloc note.	anomalous
4	"Een soort zure komkommer staat ook wel bekend als een augurk." A sort of sour cucumber is also known as a pickle.	expected
	"Een soort zure komkommer staat ook wel bekend als een lamel." A sort of sour cucumber is also known as a slat.	anomalous
5	"De metselaar zette op de laag nat cement weer een baksteen." The (brick)layer put on the wet layer of cement another brick.	expected
	"De metselaar zette op de laag nat cement weer een slaapzak." The (brick)layer put on the wet layer of cement another sleeping bag.	anomalous
6	"De aap had honger en pelde een banaan." The monkey was hungry and peeled a banana.	expected
	"De aap had honger en pelde een kegel." The monkey was hungry and peeled a pin.	anomalous
7	"In het bekende liedje speelt Mien op een mandoline terwijl Jo plukt aan de snaren van een banjo." In the famous song Mien plays the mandolin while Jo plucks the strings of a banjo.	expected
	"In het bekende liedje speelt Mien op een mandoline terwijl Jo plukt aan de snaren van een pipet." In the famous song Mien plays the mandolin while Jo plucks the strings of a pipette.	anomalous
8	"Michael Jordan dribbelde tussen de verdedigers door en gooide de bal precies in de basket." Michael Jordan dribbled through the defenders and threw the ball right in the basket.	expected
	"Michael Jordan dribbelde tussen de verdedigers door en gooide de bal precies in de vinvis." Michael Jordan dribbled through the defenders and threw the ball right in the whale.	anomalous
9	"Knie, heup, enkel en voet zijn allemaal onderdeel van een lichaamsdeel genaamd been." Knee, hip, ankle and foot are all part of a body part called leg.	expected
	"Knie, heup, enkel en voet zijn allemaal onderdeel van een lichaamsdeel genaamd ding." Knee, hip, ankle and foot are all part of a body part called thing.	anomalous
10	"Het houten voorwerp dat de jagende aboriginal naar zijn prooi slingerde kwam terug want het was een boemerang." The wooden object that the hunting aboriginal hurled at his prey, returned because it was a boomerang.	expected
	"Het houten voorwerp dat de jagende aboriginal naar zijn prooi slingerde kwam terug want het was een eierdop." The wooden object that the hunting aboriginal hurled at his prey, returned because it was an egg cup.	anomalous

APPENDIX

Nr.	sentence	Condition
11	"Op de ambassade van het vijandige land ontplofte een bom." At the embassy of the hostile country exploded a bomb.	expected
	"Op de ambassade van het vijandige land ontplofte een reus." At the embassy of the hostile country exploded a giant.	anomalous
12	"De kok schepte het warme eten op een bord." The chef scooped the hot food on a plate.	expected
	"De kok schepte het warme eten op een veld." The chef scooped the hot food on a field.	anomalous
13	"Voordat e-mail bestond stuurden verliefde stelletjes elkaar soms via de post een brief." Before the e-mail existed, love birds would send each other sometimes through the mail a letter.	expected
	"Voordat e-mail bestond stuurden verliefde stelletjes elkaar soms via de post een trap." Before the e-mail existed, love birds would send each other sometimes through the mail a stair.	anomalous
14	"Voor in de auto brandde Piet wat mp3'tjes op een cd." For in the car, Piet burned some mp3's on a cd.	expected
	"Voor in de auto brandde Piet wat mp3'tjes op een fauteuil." For in the car, Piet burned some mp3's on an armchair.	anomalous
15	"Haar favoriete merk was Verkade en dan die met de roze wikkel als het aankwam op het eten van een chocoladereep." Her favorite brand was Verkade and then the one with the pink wrap when it came to eating chocolate bars.	expected
	"Haar favoriete merk was Verkade en dan die met de roze wikkel als het aankwam op het eten van een pantserwagen." Her favorite brand was Verkade and then the one with the pink wrap when it came to eating armored cars.	anomalous
16	"Met een vertrokken gezicht beet het jongetje in de zure gele vrucht genaamd citroen." While grimacing the boy bit in to the yellow fruit called lemon.	expected
	"Met een vertrokken gezicht beet het jongetje in de zure gele vrucht genaamd matras." While grimacing the boy bit in to the yellow fruit called mattress	anomalous
17	"Om de kamer binnen te komen draaide Jan aan de klink van een deur." To enter the room, Jan twisted the doorknob of a door.	expected
	"Om de kamer binnen te komen draaide Jan aan de klink van een God." To enter the room, Jan twisted the doorknob of a God.	anomalous
18	"Met een atletische draaibeweging die de Grieken vroeger al gebruikten wierp Rutger Smith een discus." With an athletic turning movement that was already used by the ancient Greeks, Rutger Smith threw a discus.	expected

Nr.	sentence	Condition
	"Met een atletische draaibeweging die de Grieken vroeger al gebruikten wierp Rutger Smith een vriezer." With an athletic turning movement that was already used by the ancient Greeks, Rutger Smith threw a freezer.	anomalous
19	"Het dikste en kortste uitsteeksel aan je hand, met maar twee kootjes, is je duim." The fattest and shortest protrusion of your hand, with only two phalanges, is your thumb.	expected
	"Het dikste en kortste uitsteeksel aan je hand, met maar twee kootjes, is je chef." The fattest and shortest protrusion of your hand, with only two phalanges, is your chef.	anomalous
20	"Tegenwoordig bekijkt men thuis films in goede kwaliteit op een schijfje genaamd dvd." People watch movies these days at home in good quality on a disc called dvd.	expected
	"Tegenwoordig bekijkt men thuis films in goede kwaliteit op een schijfje genaamd omelet." People watch movies these days at home in good quality on a disc called omelet.	anomalous
21	"De notaris plakte een postzegel op de buitenkant van een envelop." The notary attached a stamp on the outer side of an envelope.	expected
	"De notaris plakte een postzegel op de buitenkant van een bewaker." The notary attached a stamp on the outer side of a guard.	anomalous
22	"Om ruimte te besparen in de stad bouwde de gemeente hoger en verving een deel van de laagbouw door een flat." To save space in the city, the council built higher and replaced a part of the low-rise by a flat.	expected
	"Om ruimte te besparen in de stad bouwde de gemeente hoger en verving een deel van de laagbouw door een kern." To save space in the city, the council built higher and replaced a part of the low-rise by a core.	anomalous
23	"De Rattenvanger van Hamelen lokte de ratten met zich mee door het bespelen van een fluit." The rat catcher from Hamelen lured the rats to him by playing a fluit.	expected
	"De Rattenvanger van Hamelen lokte de ratten met zich mee door het bespelen van een klep." The rat catcher from Hamelen lured the rats to him by playing a valve.	anomalous
24	"Kinderen gooien 's zomers op het stand vaak over met een draaiende plastic frisbee." In summer children often throw around on the beach a turning plastic frisbee.	expected
	"Kinderen gooien 's zomers op het stand vaak over met een draaiende plastic wasbak." In summer children often throw around on the beach a turning plastic sink.	anomalous

Nr.	sentence	Condition
25	"Jimi Hendrix werd bekend door zijn virtuoze spel op zijn gitaar." Jimi Hendrix became famous because of his virtuoso on his guitar.	expected
	"Jimi Hendrix werd bekend door zijn virtuoze spel op zijn barak." Jimi Hendrix became famous because of his virtuoso on his barrack.	anomalous
26	"Tiger Woods wilde een hole-in-one slaan en haalde flink uit met zijn golfclub." Tiger Woods was aiming for a hole-in-one and vigorously swung his golf club.	expected
	"Tiger Woods wilde een hole-in-one slaan en haalde flink uit met zijn eindstand." Tiger Woods was aiming for a hole-in-one and vigorously swung his end score.	anomalous
27	"Nadat de chocoladeliefhebber haar boterham had besmeerd met boter bestrooide ze hem met flink veel hagelslag." After the chocolate lover buttered her sandwich, she sprinkled on it a lot of chocolate sprinkles.	expected
	"Nadat de chocoladeliefhebber haar boterham had besmeerd met boter bestrooide ze hem met flink veel projector." After the chocolate lover buttered her sandwich, she sprinkled on it a lot of projectors.	anomalous
28	"De timmerman sloeg per ongeluk keihard op zijn hand met een hamer." The handyman accidently hit his hand really hard with a hammer.	expected
	"De timmerman sloeg per ongeluk keihard op zijn hand met een leuning." The handyman accidently hit his hand really hard with a handrail.	anomalous
29	"Een keer was een soldaat voordat hij wierp vergeten het pinnetje te verwijderen uit een handgranaat." One time a soldier before throwing (he) forgot to remove the pin from a hand grenade.	expected
	"Een keer was een soldaat voordat hij wierp vergeten het pinnetje te verwijderen uit een plankenvloer." One time a soldier before throwing (he) forgot to remove the pin from a hardwood floor.	anomalous
30	"Omdat het weiland met sloten omgeven was stond alleen op de dam een houten hek." Because the meadow was surrounded by creaks, only on the dam there was a wooden gate.	expected
	"Omdat het weiland met sloten omgeven was stond alleen op de dam een houten mes." Because the meadow was surrounded by creaks, only on the dam there was a wooden knife.	anomalous
31	"Zelfs zonder hard nadenken worden er telkens stroompjes doorgegeven tussen neuronen in je hersenen." Even without thinking hard, signals are constantly passed on between neurons in your brain.	expected

Nr.	sentence	Condition
	"Zelfs zonder hard nadenken worden er telkens stroompjes doorgegeven tussen neuronen in je boerderij." Even without thinking hard, signals are constantly passed on between neurons in your farm.	anomalous
32	"De hockeyspeler sloeg de bal naar voren met zijn hockeystick." The hockey player hit the ball to the front with his hockey stick.	expected
	"De hockeyspeler sloeg de bal naar voren met zijn paardekop." The hockey player hit the ball to the front with his horse head.	anomalous
33	"De ouderwetse Engelsman droeg een paraplu en om zijn kaalheid te verbergen ook nog een hoed." The old-fashioned Englishman wore an umbrella and to hide his boldness also a hat.	expected
	"De ouderwetse Engelsman droeg een paraplu en om zijn kaalheid te verbergen ook nog een wang." The old-fashioned Englishman wore an umbrella and to hide his boldness also a cheek.	anomalous
34	"Op het schoolplein draaide een meisje rond haar middel zwierig met een hoepel." At the school yard a girl whirled her waist elegantly with a hula hoop.	expected
	"Op het schoolplein draaide een meisje rond haar middel zwierig met een cocon." At the school yard a girl whirled her waist elegantly with a cocoon.	anomalous
35	"Omdat hij binnen vijf minuten weg wilde keek Bert telkens op zijn horloge." Because he wanted to leave in five minutes, Bert constantly checked his watch.	expected
	"Omdat hij binnen vijf minuten weg wilde keek Bert telkens op zijn kapitaal." Because he wanted to leave in five minutes, Bert constantly checked his capital.	anomalous
36	"Wanneer de stekker er niet in zit kan een muzikant wel de toetsen indrukken maar komt er geen geluid uit zijn keyboard." When the plug is not plugged in, a musician can try and push buttons, but no sound will come from his keyboard.	expected
	"Wanneer de stekker er niet in zit kan een muzikant wel de toetsen indrukken maar komt er geen geluid uit zijn skipak." When the plug is not plugged in, a musician can try and push buttons, but no sound will come from his ski outfit.	anomalous
37	"Op de tegels met putjes rolde het spelende schoolkind met een prachtige glazen knikker." On the tiles with drains a school child played with a beautiful glass marble.	expected
	"Op de tegels met putjes rolde het spelende schoolkind met een prachtige glazen luipaard." On the tiles with drains a school child played with a beautiful glass leopard.	anomalous

Nr.	sentence	Condition
38	"Bij de thee eten we altijd een koekje." With thee we always eat a cookie.	expected
	"Bij de thee eten we altijd een muisje." With thee we always eat a mouse.	anomalous
39	"Tijdens het navigeren wist de padvinder precies waar het noorden was wanneer hij een blik wierp op zijn kompas." While navigating the boy scout always knew exactly where to find north when checking his compass.	expected
	"Tijdens het navigeren wist de padvinder precies waar het noorden was wanneer hij een blik wierp op zijn budget." While navigating the boy scout always knew exactly where to find north when checking his budget.	anomalous
40	"Het kind wilde een ketting maken en reeg daarom aan het touw telkens opnieuw een houten kraal." The child wanted to make a necklace and therefore strung to a cord over and over again a wooden bead.	expected
	"Het kind wilde een ketting maken en reeg daarom aan het touw telkens opnieuw een houten flap." The child wanted to make a necklace and therefore strung to a cord over and over again a wooden flap.	anomalous
41	"Om de dakgoot te bereiken moest de schilder omhoog klimmen langs de sporten van een ladder." To reach the gutter, the painter had to climb the steps of the ladder.	expected
	"Om de dakgoot te bereiken moest de schilder omhoog klimmen langs de sporten van een borrel." To reach the gutter, the painter had to climb the steps of a drink.	anomalous
42	"Het donkere steegje achter mijn huis wordt 's avonds verlicht door een lantaarnpaal." The dark alley behind my house becomes illuminated at night by a lamppost.	expected
	"Het donkere steegje achter mijn huis wordt 's avonds verlicht door een schommelstoel." The dark alley behind my house becomes illuminated at night by a rocking chair.	anomalous
43	"Opa zette de naald van de platenspeler op de ouderwetse LP." Grandpa adjusted the needle of the record player on the old-fashioned LP.	expected
	"Opa zette de naald van de platenspeler op de ouderwetse anijs." Grandpa adjusted the needle of the record player on the old-fashioned anise.	anomalous
44	"In 1969 zette Neil Armstrong als eerste mens voet op de maan." In 1969 Neil Armstrong was the first person to ever set foot upon the moon.	expected
	"In 1969 zette Neil Armstrong als eerste mens voet op de boot." In 1969 Neil Armstrong was the first person to ever set foot upon the boat.	anomalous

Nr.	sentence	Condition
45	"Aan de achterkant van een bout bevestig je vaak een metalen moer." On the back of a bolt you often secure a metal nut.	expected
	"Aan de achterkant van een bout bevestig je vaak een metalen piek." On the back of a bolt you often secure a metal top.	anomalous
46	"Omdat ze zulke koude oren had in de sneeuw droeg het meisje een warme muts." Because of her cold ears due to the snow, the girl wore a warm beanie.	expected
	"Omdat ze zulke koude oren had in de sneeuw droeg het meisje een warme kruk." Because of her cold ears due to the snow, the girl wore a warm stool.	anomalous
47	"De agente die de automobilist door rood zag rijden noteerde snel de letters en getallen van zijn nummerbord." The police officer, who saw the car driver hitting a red light, quickly wrote down the letters and numbers of his number plate.	expected
	"De agente die de automobilist door rood zag rijden noteerde snel de letters en getallen van zijn korenveld." The police officer, who saw the car driver hitting a red light, quickly wrote down the letters and numbers of his cornfield.	anomalous
48	"Katrien druppelde vloeistof op de contactlens en deed hem voorzichtig in haar oog." Katrien drippled liquid on the contact lens and carefully put it in her eye.	expected
	"Katrien druppelde vloeistof op de contactlens en deed hem voorzichtig in haar hooi." Katrien drippled liquid on the contact lens and carefully put it in her hay.	anomalous
49	"Sven wilde het liefst stroop én poedersuiker op zijn pannenkoek." Sven preferred both syrup and icing sugar on his pancake.	expected
	"Sven wilde het liefst stroop én poedersuiker op zijn locatie." Sven preferred both syrup and icing sugar on his location.	anomalous
50	"De duikers waren op zoek naar sieraden en haalden uit een geopende oester een kleine parel." The divers were looking for jewelry and removed from an opened oyster a small pearl.	expected
	"De duikers waren op zoek naar sieraden en haalden uit een geopende oester een kleine fakkel." The divers were looking for jewelry and removed from an opened oyster a small torch.	anomalous
51	"Tussen alle andere auto's bij het pretpark kon de automobilist pas na lang zoeken zijn auto kwijt op een parkeerplaats." Due to all the other cars at the amusement park, it took the car driver a long time to find a spot on the parking lot.	expected
	"Tussen alle andere auto's bij het pretpark kon de automobilist pas na lang zoeken zijn auto kwijt op een brandstapel." Due to all the other cars at the amusement park, it took the car driver a long time to find a spot on the pyre.	anomalous

Nr.	sentence	Condition
52	"De slagwerker in het klassieke orkest produceerde een zwaar en diep geluid met een pauk." The percussionist in the classical orchestra produced a deep and heavy sound with a kettledrum.	expected
	"De slagwerker in het klassieke orkest produceerde een zwaar en diep geluid met een mijt." The percussionist in the classical orchestra produced a deep and heavy sound with a mite.	anomalous
53	"Om op de rivier vooruit te komen in zijn kano gebruikte de indiaan een houten peddel." In order to float forward on the river in his canoe, the Native American used a wooden paddle.	expected
	"Om op de rivier vooruit te komen in zijn kano gebruikte de indiaan een houten sesam." In order to float forward on the river in his canoe, the Native American used wooden sesame.	anomalous
54	"Om met blauwe inkt haar handtekening te kunnen zetten pakte de directrice haar pen." In order to write her signature with blue ink, the principal grabbed her pen.	expected
	"Om met blauwe inkt haar handtekening te kunnen zetten pakte de directrice haar zeep." In order to write her signature with blue ink, the principal grabbed her soap.	anomalous
55	"De Italiaanse kok zette de oven alvast aan en deed tomaten, kaas en champignons op een pizza." The Italian chef turned on the oven already and put the tomatoes, cheese and mushrooms on a pizza.	expected
	"De Italiaanse kok zette de oven alvast aan en deed tomaten, kaas en champignons op een shampoo." The Italian chef turned on the oven already and put the tomatoes, cheese and mushrooms on a shampoo.	anomalous
56	"De gitarist ramde op de snaren met een stukje plastic genaamd plectrum." The guitarist hammered on the strings with a piece of plastic called plectrum.	expected
	"De gitarist ramde op de snaren met een stukje plastic genaamd inktstel." The guitarist hammered on the strings with a piece of plastic called inkstand.	anomalous
57	"De beste atleten kunnen over een zes meter hoge balk springen, waarbij ze gebruikmaken van een polsstok." The best athletes can jump over a six meter high beam using a jumping pole.	expected
	"De beste atleten kunnen over een zes meter hoge balk springen, waarbij ze gebruikmaken van een sproeier." The best athletes can jump over a six meter high beam using a sprinkler.	anomalous

Nr.	sentence	Condition
58	"Met een puntenslijper sleep de tekenaar de punt van zijn potlood." With a pencil sharpener the artist sharpened the point of his pencil.	expected
	"Met een puntenslijper sleep de tekenaar de punt van zijn bestek." With a pencil sharpener the artist sharpened the point of his cutlery.	anomalous
59	"De groen-witte groente die uit dezelfde familie komt als sjalotjes en knoflook heet prei." The green-white vegetable from the same family as the shallot and the garlic is called leek.	expected
	"De groen-witte groente die uit dezelfde familie komt als sjalotjes en knoflook heet kous." The green-white vegetable from the same family as the shallot and the garlic is called stocking.	anomalous
60	"Bert prikte het geboortekaartje met een punaise naast de kerstkaarten op zijn prikbord." Bert pinned the birth card with a drawing pin next to the Christmas cards on his bulletin board.	expected
	"Bert prikte het geboortekaartje met een punaise naast de kerstkaarten op zijn bloemperk." Bert pinned the birth card with a drawing pin next to the Christmas cards on his flower bed.	anomalous
61	"Om in het riool onder de straat te komen opende de brandweerman een put." To enter the sewer beneath the street, the fireman opened a drain.	expected
	"Om in het riool onder de straat te komen opende de brandweerman een kool." To enter the sewer beneath the street, the fireman opened a cabbage.	anomalous
62	"De oude man zat bij de vensterbank en keek naar buiten door zijn raam." The old man sat on the windowsill and looked outside through his window.	expected
	"De oude man zat bij de vensterbank en keek naar buiten door zijn dier." The old man sat on the windowsill and looked outside through his animal.	anomalous
63	"Zijn broek zat te wijd dus trok hij aan zijn riem." His pants were too loose so he adjusted his belt.	expected
	"Zijn broek zat te wijd dus trok hij aan zijn poot." His pants were too loose so he adjusted his paw.	anomalous
64	"De Chinees gebruikte altijd eetstokjes bij het eten van zijn witte rijst." The Chinese man always used chopsticks while eating white rice.	expected
	"De Chinees gebruikte altijd eetstokjes bij het eten van zijn witte kruin." The Chinese man always used chopsticks while eating white crowns.	anomalous
65	"Bij zijn huwelijksaanzoek opende de man een klein doosje met daarin een gouden ring." During his marriage proposal, the man opened a small box with a golden ring (in it).	expected

Nr.	sentence	Condition
	"Bij zijn huwelijksaanzoek opende de man een klein doosje met daarin een gouden neef." During his marriage proposal, the man opened a small box with a golden cousin (in it).	anomalous
66	"Je moet goed in de gaten houden welke afslag je moet nemen tijdens het rondrijden over een rotonde." You should watch carefully which exit you should take when driving on a roundabout.	expected
	"Je moet goed in de gaten houden welke afslag je moet nemen tijdens het rondrijden over een stomerij." You should watch carefully which exit you should take when driving on a dry cleaner's.	anomalous
67	"Bij harde regen tijdens het autorijden gebruik ik altijd de snelste stand van mijn ruitenwisser." When it is hozing while I'm driving, I always set the fastest mode of my wipers.	expected
	"Bij harde regen tijdens het autorijden gebruik ik altijd de snelste stand van mijn drogisterij." When it is hozing while I'm driving, I always set the fastest mode of my drug store.	anomalous
68	"De jongen dacht dat het een gewoon zwaard was maar er zat een kromming in het lemmet dus was het een sabel." The boy thought it was a normal sword but there was a curve in the blade so it was a sabre.	expected
	"De jongen dacht dat het een gewoon zwaard was maar er zat een kromming in het lemmet dus was het een boetiek." The boy thought it was a normal sword but there was a curve in the blade so it was a boutique.	anomalous
69	"Candy Dulfer blies een geweldige jazzy solo op haar saxofoon." Candy Dulfer played an amazing jazzy solo on her saxophone.	expected
	"Candy Dulfer blies een geweldige jazzy solo op haar houtvezel." Candy Dulfer played an amazing jazzy solo on her woodfibre.	anomalous
70	"Sommige schilders doen maanden over het maken van een schilderij." Some painters take months to finish a painting.	expected
	"Sommige schilders doen maanden over het maken van een instituut." Some painters take months to finish an institute.	anomalous
71	"Het hebben van een bolvormig schild en een langzame tred kenmerkt de schildpad." Having a bulged shield and a slow walking pace is what characterizes the tortoise.	expected
	"Het hebben van een bolvormig schild en een langzame tred kenmerkt de dansvloer." Having a bulged shield and a slow walking pace is what characterizes the dancefloor.	anomalous

Nr.	sentence	Condition
72	"Na het diner veegde de deftige man zijn mond schoon met een servet." After dinner, the posh man wiped his mouth with a napkin.	expected
	"Na het diner veegde de deftige man zijn mond schoon met een insekt." After dinner, the posh man wiped his mouth with an insect.	anomalous
73	"De keukenhulp verkreeg jus d'orange door het persen van een stuk fruit genaamd sinaasappel." The kitchen help made orange juice by squeezing a piece of fruit called orange.	expected
	"De keukenhulp verkreeg jus d'orange door het persen van een stuk fruit genaamd portemonnee." The kitchen help made orange juice by squeezing a piece of fruit called wallet.	anomalous
74	"Toen de trein voorbij was moesten de auto's nog lang wachten op het opengaan van een slagboom." After the train had passed by, the cars still had to wait a long time for the opening of the barrier.	expected
	"Toen de trein voorbij was moesten de auto's nog lang wachten op het opengaan van een trompet." After the train had passed by, the cars still had to wait a long time for the opening of the trumpet.	anomalous
75	"Het slijmspoor op de aangevreten sla in de moestuin werd veroorzaakt door een slak." The mucus trail on the half-eaten lettuce in the kitchen garden was caused by a snail.	expected
	"Het slijmspoor op de aangevreten sla in de moestuin werd veroorzaakt door een broche." The mucus trail on the half-eaten lettuce in the kitchen garden was caused by a pin.	anomalous
76	"In het gevaarlijke oerwoud klonk het gesis van een slang." In the dangerous jungle, resounded the hissing of a snake.	expected
	"In het gevaarlijke oerwoud klonk het gesis van een baard." In the dangerous jungle, resounded the hissing of a beard.	anomalous
77	"Het kokshulpje wilde een courgette snijden en legde hem klaar op een snijplank." The cook's boy wanted to chop a zucchini and arranged it on a chopping board.	expected
	"Het kokshulpje wilde een courgette snijden en legde hem klaar op een prijslijst." The cook's boy wanted to chop a zucchini and arranged it on a price list.	anomalous
78	"De mammoetjager zag een mammoet, kwam uit zijn hinderlaag en wierp zijn houten speer." The mammoth hunter saw a mammoth, appeared from his ambush and threw his wooden spear.	expected

Nr.	sentence	Condition
	"De mammoetjager zag een mammoet, kwam uit zijn hinderlaag en wierp zijn houten doorn." The mammoth hunter saw a mammoth, appeared from his ambush and threw his wooden thorn.	anomalous
79	"Als vrouwen zich opmaken voor een avondje uit, staan ze vaak uren voor hun spiegel." When women get ready for a night out, they often spend hours in front of the mirror.	expected
	"Als vrouwen zich opmaken voor een avondje uit, staan ze vaak uren voor hun dichter." When women get ready for a night out, they often spend hours in front of the poet.	anomalous
80	"Timmeren was niet Jeroens specialiteit want zijn scheurkalender hing aan een slordig in de muur geslagen spijker." Carpentry was not Jeroen's specialty because his tear-off calendar was pinned to the wall with a poorly hammered nail.	expected
	"Timmeren was niet Jeroens specialiteit want zijn scheurkalender hing aan een slordig in de muur geslagen portiek." Carpentry was not Jeroen's specialty because his tear-off calendar was pinned to the wall with a poorly hammered porch.	anomalous
81	"De trein rijdt doorgaans over een spoor." The train usually runs over a trail.	expected
	"De trein rijdt doorgaans over een blad." The train usually runs over a leave.	anomalous
82	"De Fransman nam een slok wijn en smeerde kruidenboter op zijn vers gesneden stokbrood." The Frenchman took a sip of wine and spread some herb butter on his freshly cut baguette.	expected
	"De Fransman nam een slok wijn en smeerde kruidenboter op zijn vers gesneden brandhout." The Frenchman took a sip of wine and spread some herb butter on his freshly cut fire wood.	anomalous
83	"Het verkeer bij de kruising stond te wachten voor een stoplicht." The traffic by the crossing was waiting for the traffic lights.	expected
	"Het verkeer bij de kruising stond te wachten voor een borstbeen." The traffic by the crossing was waiting for the breastbone.	anomalous
84	"Ondanks dat de man had gedronken stapte hij toch in de auto en kroop achter zijn stuur." Even though the man had been drinking, he stepped in the car behind his wheel.	expected
	"Ondanks dat de man had gedronken stapte hij toch in de auto en kroop achter zijn blok." Even though the man had been drinking, he stepped in the car behind his block.	anomalous

Nr.	sentence	Condition
85	"Het jarige buurjongetje blies zeven kaarsjes uit op een door zijn moeder gebakken taart." The birthday boy from next door blew out seven candles on made by his mother a cake.	expected
	"Het jarige buurjongetje blies zeven kaarsjes uit op een door zijn moeder gebakken vloot." The birthday boy from next door blew out seven candles on a made by his mother fleet.	anomalous
86	"Toen er eindelijk een gezin het restaurant in kwam dekte de ober gauw een tafel." When a family finally entered the restaurant, the waiter quickly set a table.	expected
	"Toen er eindelijk een gezin het restaurant in kwam dekte de ober gauw een kennis." When a family finally entered the restaurant, the waiter quickly set an acquaintance.	anomalous
87	"Op de camping begon de campeerder met het opzetten van zijn tent." On the camping, the camping guest started to put up his tent.	expected
	"Op de camping begon de campeerder met het opzetten van zijn hoorn." On the camping, the camping guest started to put up his horn.	anomalous
88	"Pastasaus krijgt zijn rode kleur door een groente genaamd tomaat." The pasta sauce is red because of a vegetable called tomato.	expected
	"Pastasaus krijgt zijn rode kleur door een groente genaamd beugel." The pasta sauce is red because of vegetable called braces.	anomalous
89	"Voor een klokjeschtig geluid sloeg de amuzikale Henk met een klein metalen staafje voorzichtig op een triangel." To get a bell-like sound, the tone-deaf Henk hit carefully with a metal stick a triangle.	expected
	"Voor een klokjeschtig geluid sloeg de amuzikale Henk met een klein metalen staafje voorzichtig op een sloppenwijk." To get a bell-like sound, the tone-deaf Henk hit carefully with a metal stick a slum area.	anomalous
90	"Jans ogen begonnen te tranen door het snijden van een ui." Jan's eyes started to tear up when cutting an onion.	expected
	"Jans ogen begonnen te tranen door het snijden van een baai." Jan's eyes started to tear up when cutting a bay.	anomalous
91	"Sommige videotheken verhuren nog steeds af en toe een film op videoband." Some video stores still rent out once in a while films on videotapes.	expected
	"Sommige videotheken verhuren nog steeds af en toe een film op studiegenoot." Some video stores still rent out once in a while films on study partners.	anomalous
92	"Mark maakte het bekende boze handgebaar door het strekken van zijn middelste vinger." Mark made the famous angry hand gesture by stretching his middle finger.	expected

Nr.	sentence	Condition
	"Mark maakte het bekende boze handgebaar door het strekken van zijn middelste lezer." Mark made the famous angry hand gesture by stretching his middle reader.	anomalous
93	"Patrick Kluivert nam een penalty en schopte met alle kracht die hij had tegen de voetbal." Patrick Kluivert took a penalty and kicked with all his strength the football.	expected
	"Patrick Kluivert nam een penalty en schopte met alle kracht die hij had tegen de biceps." Patrick Kluivert took a penalty and kicked with all his strength the biceps.	anomalous
94	"De boer prakte zijn maaltijd altijd met zijn vork." The farmer always mashed his meal with his fork.	expected
	"De boer prakte zijn maaltijd altijd met zijn sloot." The farmer always mashed his meal with his ditch.	anomalous
95	"De agressieve jongen wilde vechten en balde zijn rechter vuist." The aggressive boy wanted to fight and clenched his right fist.	expected
	"De agressieve jongen wilde vechten en balde zijn rechter bocht." The aggressive boy wanted to fight and clenched his right corner.	anomalous
96	"Omdat iemand tegen mijn fiets had geschopt zat er een slag in mijn wiel." Because someone kicked my bike, it now has a buckled wheel.	expected
	"Omdat iemand tegen mijn fiets had geschopt zat er een slag in mijn deeg." Because someone kicked my bike, it now has a buckled dough.	anomalous
97	"De merel voerde haar jongen in het nest een door haarzelf uit de grond getrokken worm." The blackbird fed her young ones in the nest a herself pulled out of the ground worm.	expected
	"De merel voerde haar jongen in het nest een door haarzelf uit de grond getrokken sprei." The blackbird fed her young ones in the nest a herself pulled out of the ground bedspread.	anomalous
98	"Hutspot krijgt zijn kleur van een oranje groente genaamd wortel." 'Hutspot' is orange because of the vegetable called carrot.	expected
	"Hutspot krijgt zijn kleur van een oranje groente genaamd cursus." 'Hutspot' is orange because of the vegetable called course.	anomalous
99	"Het was warm want gedurende de hele dag scheen de zon." It was warm because the whole day was shining the sun.	expected
	"Het was warm want gedurende de hele dag scheen de vis." It was warm because the whole day was shining the fish.	anomalous
100	"Het schip werd stevig vastgelegd met een anker" The ship was firmly secured with an anchor.	expected
	"Het schip werd stevig vastgelegd met een rijtuig" The ship was firmly secured with a vehicle.	anomalous
101	"Ze kochten een wieg voor de baby" They bought a crib for the baby.	expected

Nr.	sentence	Condition
	"Ze kochten een wieg voor de visie" They bought a crib for the vision.	anomalous
102	"Voor de ontspanning neem ik graag een warm bad" To relax I like taking a warm bath.	expected
	"Voor de ontspanning neem ik graag een warm loon" To relax I like taking a warm pay check.	anomalous
103	"In de lente zet ik bloembakken op het balkon" In spring I put flower boxes on the balcony.	expected
	"In de lente zet ik bloembakken op het gerecht" In spring I put flower boxes on the meal.	anomalous
104	"Voor het kinderfeest hebben we al een opgeblazen ballon" For the kids' party we already have a blown up balloon.	expected
	"Voor het kinderfeest hebben we al een opgeblazen puzzel" For the kids' party we already have a blown up puzzle.	anomalous
105	"Haar walkman deed het niet meer vanwege de batterij" Her walkman didn't work anymore because of the battery.	expected
	"Haar walkman deed het niet meer vanwege de ananas" Her walkman didn't work anymore because of the ananas.	anomalous
106	"Het team speelde goed en won natuurlijk de beker" The team played well and naturally won the cup.	expected
	"Het team speelde goed en won natuurlijk de lever" The team played well and naturally won the liver.	anomalous
107	"Hij veegde de vloer met een bezem" He wiped the floor with a broom.	expected
	"Hij veegde de vloer met een sukkel" He wiped the floor with an idiot.	anomalous
108	"Hij hakte de boom om met een bijl" He chopped the tree with an axe.	expected
	"Hij hakte de boom om met een map" He chopped the tree with a folder.	anomalous
109	"In de wei plukte het meisje een bloem" In the meadow the girl plucked a flower.	expected
	"In de wei plukte het meisje een wolk" In the meadow the girl plucked a cloud.	anomalous
110	"Mijn opa kan niets lezen zonder zijn bril" My grandpa can't read without his glasses.	expected
	"Mijn opa kan niets lezen zonder zijn tocht" My grandpa can't read without his breeze.	anomalous
111	"Ze kunnen de rivier niet over vanwege een reparatie aan de brug" They can't cross the river due to a renovation of the bridge.	expected
	"Ze kunnen de rivier niet over vanwege een reparatie aan de moord" They can't cross the river due to a renovation of the murder.	anomalous
112	"De chauffeur stempelde mijn strippenkaart in de bus" The driver stamped my bus ticket in the bus.	expected

Nr.	sentence	Condition
	"De chauffeur stempelde mijn strippenkaart in de hut" The driver stamped my bus ticket in the hut.	anomalous
113	"Haar moeder is ziek maar wil niet naar de dokter" Her mother is ill but she doesn't want to go to the doctor.	expected
	"Haar moeder is ziek maar wil niet naar de invloed" Her mother is ill but she doesn't want to go to the influence.	anomalous
114	"Wijn wordt gemaakt van druiven" Wine is made of grapes.	expected
	"Wijn wordt gemaakt van paters" Wine is made of priests.	anomalous
115	"In het park voerden de kinderen de eend" In the park the children fed the duck.	expected
	"In het park voerden de kinderen de roos" In the park the children fed the rose.	anomalous
116	"De schoonmaakster gooit water uit een emmer" The cleaning lady throws water out of a bucket.	expected
	"De schoonmaakster gooit water uit een sofa" The cleaning lady throws water out of a sofa.	anomalous
117	"In die fabriek bottelt men wijn in verschillende soorten flessen" In that factory, wine is bottled in different kind of bottles.	expected
	"In die fabriek bottelt men wijn in verschillende soorten woningen" In that factory, wine is bottled in different kind of homes.	anomalous
118	"Om plantjes water te geven gebruik je een gieter" To water the plants you use a watering can.	expected
	"Om plantjes water te geven gebruik je een pijler" To water the plants you use a pillar.	anomalous
119	"Om de kamer donker te maken, heb ik zwarte gordijnen" To make the room dark, I have black curtains.	expected
	"Om de kamer donker te maken, heb ik zwarte momenten" To make the room dark, I have black moments.	anomalous
120	"De vijand van Peter Pan heeft geen hand maar een haak" The enemy of Peter Pan doesn't have a hand but a hook.	expected
	"De vijand van Peter Pan heeft geen hand maar een douche" The enemy of Peter Pan doesn't have a hand but a shower.	anomalous
121	"De valentijnskaart had de vorm van een hart" The Valentine's card had the shape of a heart.	expected
	"De valentijnskaart had de vorm van een beeld" The Valentine's card had the shape of a statue.	anomalous
122	"Het huis van een eskimo is een iglo" The house of an eskimo is an iglo.	expected
	"Het huis van een eskimo is een file" The house of an eskimo is a traffic jam.	anomalous
123	"Mijn favoriete zomerdessert is natuurlijk een ijsje" My favorite summer dessert is of course an ice cream.	expected

Nr.	sentence	Condition
	"Mijn favoriete zomerdessert is natuurlijk een veulen" My favorite summer dessert is of course a foal.	anomalous
124	"Ik hou van kerst want ik krijg altijd een kado" I love Christmas because I always get a gift.	expected
	"Ik hou van kerst want ik krijg altijd een aapje" I love Christmas because I always get a monkey.	anomalous
125	"Vanwege ontstekingen trok de tandarts haar kies" Because of an inflammation the dentist pulled out her molar.	expected
	"Vanwege ontstekingen trok de tandarts haar duin" Because of an inflammation the dentist pulled out her dune.	anomalous
126	"Kermit is een heel sympathieke kikker" Kermit is a very sympathetic frog.	expected
	"Kermit is een heel sympathieke bedding" Kermit is a very sympathetic bedding.	anomalous
127	"Carla bewaart haar oude poppen in een houten kist" Carla keeps her old dolls in a wooden chest.	expected
	"Carla bewaart haar oude poppen in een houten poort" Carla keeps her old dolls in a wooden gate.	anomalous
128	"De boer melkte de koe" The farmer milks the cow.	expected
	"De boer melkte de mie" The farmer milks the Chinese noodles.	anomalous
129	"Mijn konijn zit nooit opgesloten in de kooi" My rabbit is never locked up inside the cage.	expected
	"Mijn konijn zit nooit opgesloten in de soep" My rabbit is never locked up inside the soup.	anomalous
130	"De wijnfles zat nog dicht met een kurk" The wine bottle was still sealed with a cork.	expected
	"De wijnfles zat nog dicht met een vonk" The wine bottle was still sealed with a spark.	anomalous
131	"Aan de logeerders geef ik liever een zacht kussen" To the guests I'd rather give a soft pillow.	expected
	"Aan de logeerders geef ik liever een zacht thema" To the guests I'd rather give a soft theme.	anomalous
132	"Voor de verlichting naast mijn bed heb ik een lamp" To light up the room I have next to my bed a lamp.	expected
	"Voor de verlichting naast mijn bed heb ik een gids" To light up the room I have next to my bed a guide.	anomalous
133	"Soep eet je met een lepel" You eat soup with a spoon.	expected
	"Soep eet je met een tabak" You eat soup with a tobacco.	anomalous
134	"Om een rechte lijn te trekken, gebruik ik een liniaal" To draw a straight line I use a ruler.	expected

Nr.	sentence	Condition
	"Om een rechte lijn te trekken, gebruik ik een pasfoto" To draw a straight line I use a passport photo.	anomalous
135	"We staken het vuur aan met een lucifer" We lit up the fire with a match.	expected
	"We staken het vuur aan met een militair" We lit up the fire with a military.	anomalous
136	"Het Carnaval van Venetie staat bekend om zijn maskers" The Carnival of Venice is known for its masks.	expected
	"Het Carnaval van Venetie staat bekend om zijn broodjes" The Carnival of Venice is known for its bread rolls.	anomalous
137	"In de slaapkamer zwierf een irritante mug" In the bedroom there was an annoying mosquito.	expected
	"In de slaapkamer zwierf een irritante geul" In the bedroom there was an annoying trench.	anomalous
138	"Pinoccio was een jongen met een lange neus" Pinocchio was a boy with a long nose.	expected
	"Pinoccio was een jongen met een lange hoek" Pinocchio was a boy with a long corner.	anomalous
139	"Met haar nagel trok moeder een ladder in haar panty" Mother made with her nail a ladder in her tights.	expected
	"Met haar nagel trok moeder een ladder in haar robijn" Mother made with her nail a ladder in her ruby.	anomalous
140	"Een vogel die alles herhaalt is een papegaai" A bird that repeats everything is a parrot.	expected
	"Een vogel die alles herhaalt is een caravan" A bird that repeats everything is a caravan.	anomalous
141	"Om zijn kale hoofd te beschermen, draagt hij een mooie pet" To pretect is bold head, he wore a nice cap.	expected
	"Om zijn kale hoofd te beschermen, draagt hij een mooie kip" To pretect is bold head, he wore a nice chicken.	anomalous
142	"De filosoof genoot van het roken van zijn pijp" The philosopher enjoyed smoking his pipe.	expected
	"De filosoof genoot van het roken van zijn knie" The philosopher enjoyed smoking his knee.	anomalous
143	"Voor Halloween knutselden de kinderen een grote pompoen" For Halloween the children handcrafted a large pumpkin.	expected
	"Voor Halloween knutselden de kinderen een grote hangar" For Halloween the children handcrafted a large hangar.	anomalous
144	"Als we naar de bioscoop gaan, eten we altijd popcorn" When we go to the cinema, we always eat popcorn.	expected
	"Als we naar de bioscoop gaan, eten we altijd klimtouw" When we go to the cinema, we always eat climbing rope.	anomalous
145	"De Egyptenaren bouwden een grote pyramide" The Egyptians built a big pyramid.	expected

Nr.	sentence	Condition
	"De Egyptenaren bouwden een grote alligator" The Egyptians built a big alligator.	anomalous
146	"Vanwege die flat wast de glazenwasser iedere maand de ramen" Because of that flat, the window-cleaner washes every month the windows.	expected
	"Vanwege die flat wast de glazenwasser iedere maand de feiten" Because of that flat, the window-cleaner washes every month the facts.	anomalous
147	"Ik moet dit papier knippen maar heb helaas geen schaar" I have to cut this paper but I don't have a scissor.	expected
	"Ik moet dit papier knippen maar heb helaas geen pruik" I have to cut this paper but I don't have a wig.	anomalous
148	"Om de stroom uit te zetten, drukte hij op de schakelaar" To turn off the power, he pushed the button.	expected
	"Om de stroom uit te zetten, drukte hij op de studente" To turn off the power, he pushed the student.	anomalous
149	"Voor Sinterklaas zetten de kinderen hun schoen" For Saint Nicolas, the children put down their shoe.	expected
	"Voor Sinterklaas zetten de kinderen hun kraan" For Saint Nicolas, the children put down their crane.	anomalous
150	"Deze twee planken moet je vast zetten met een schroef" You have to secure these two shelves with a screw.	expected
	"Deze twee planken moet je vast zetten met een plons" You have to secure these two shelves with a splash.	anomalous
151	"Om zijn nek te beschermen, droeg hij een sjaal" To pretect his neck, he wore a scarf.	expected
	"Om zijn nek te beschermen, droeg hij een haas" To pretect his neck, he wore a hare.	anomalous
152	"Hij kon het slot niet openen zonder de sleutel" He couldn't open the lock without the key.	expected
	"Hij kon het slot niet openen zonder de vreugde" He couldn't open the lock without the joy.	anomalous
153	"Ze bewaart haar lipstift en maskara in haar tas" She keeps her lipstick and mascara in her purse.	expected
	"Ze bewaart haar lipstift en maskara in haar keus" She keeps her lipstick and mascara in her choice.	anomalous
154	"Door het zand sneed de slipper tussen zijn tenen" Because of the sand, the flipflop was stinging his toes.	expected
	"Door het zand sneed de slipper tussen zijn longen" Because of the sand, the flipflop was stinging his lungs.	anomalous
155	"In de keukenhof fotografeerden de japanse toeristen de tulpen" In Keukenhof the Japanese tourists were taking pictures of the tulips.	expected
	"In de keukenhof fotografeerden de japanse toeristen de bazaar" In Keukenhof the Japanese tourists were taking pictures of the bazar.	anomalous

Nr.	sentence	Condition
156	"Ik hang heel graag voor de tv" I like hanging out in front of the TV.	expected
	"Ik hang heel graag voor de olie" I like hanging out in front of the oil.	anomalous
157	"In de wind wapperde buiten een vlag" In the wind outside was flapping a flag.	expected
	"In de wind wapperde buiten een schat" In the wind outside was flapping a treasure.	anomalous
158	"Het vogeltje kon niet opstijgen vanwege zijn gebroken vleugel" The bird couldn't take off because of his broken wing.	expected
	"Het vogeltje kon niet opstijgen vanwege zijn gebroken stapel" The bird couldn't take off because of his broken pile.	anomalous
159	"De bezoekers keken naar de uitbarsting van de vulkaan" The visitors watched the eruption of the vulcano.	expected
	"De bezoekers keken naar de uitbarsting van de bladzij" The visitors watched the eruption of the page.	anomalous
160	"Om vliegen te vangen maakt die spin een web" To catch flies, that spider makes a web.	expected
	"Om vliegen te vangen maakt die spin een pit" To catch flies, that spider makes a seed.	anomalous

K. Auditory target words (sentence-final-words) of chapter 5 with English translations in parenthesis.

aapje (monkey), armband (bracelet), baai (bay), baard (beard), baby (baby), baksteen (brick), banjo (banjo), batterij (battery), bedding (river bed), been (leg), bestek (cutlery), beugel (braces), bezem (broom), bloem (flower), bloemperk (flowerbed), blok (block), boetiek (clothes boutique), bom (bomb), boot (boat), brandhout (fire wood), brief (letter), broodjes (sandwiches), brug (bridge), caravan (caravan), chocoladereep (chocolate bar), cocon (cocoon), cursus (course), deeg (dough), deur (door), dier (animal), dokter (doctor), doorn (thorn), douche (shower), duim (thumb), eend (duck), eierdop (egg cup), eindstand (final score), envelop (envelope), fakkel (torch), fauteuil (arm chair), file (traffic jam), flap (flap), flessen (bottles), fluit (flute), gerecht (dish), gids (guide), gitaar (guitar), gordijnen (curtains), hagelslag (chocolate sprinkles), handgranaat (hand grenade), hart (heart), hersenen (brains), hoed (hat), hoek (corner), hooi (hay), hoop (pile), horloge (watch), hut (hut), ijsje (icecream), inktstel (inkstand), insekt (insect), instituut (institute), kegel (cone), kennis (aquaintance), kern (core), kies (molar), kist (chest), klimtouw (climbing rope), knie (knee), knikker (marble), kompas (compass), kooi (cage), kruin (crown), kruk (stool), kussen (pillow), ladder (ladder), lamel (lamella), lepel (spoon), leuning (banister), lever (liver), lezer (reader), longen (lungs), loon (wage), LP (record), lucifer (match), map (folder), matras (matrass), mes (knife), mie (noodles), mijt (mite), moer (sludge), molen (mill), mug (mosquito), muisje (mouse), nummerbord (license plate), olie (oil), omelet (omelet), paardekop (horses head), pannenkoek (pancake), panty (panty), parkeerplaats (parking lot), pasfoto (passport photo), paters (priests), peddel (paddle), pet (cap), pijler (column), pit (pit), pizza (pizza), plons (splash), polsstok (jumping pole), pompoen (pumpkin), portiek (portico), prei (leek), put (well), puzzel (puzzle), pyramide (pyramid), ramen (windows), riem (belt), rijtuig (carriage), ring (ring), ruitenwisser (windshield wiper), saxofoon (saxophone), schaar (scissors), schildpad (turtle), schoen (shoe), schommelstoel (rocking chair), sinaasappel (orange), sjaal (scarve), skipak (ski suit), slak (snail), sloot (ditch), snijplank (chopping board), sofa (sofa), spiegel (mirror), spoor (track), stomerij (dry cleaners), stoplicht (traffic light), studente (female student), taart (cake), tas (bag), tent (tent), tocht (trip), triangel (triangle), trompet (trumpet), tulpen (tulips), veld (field), videoband (video tape), vinvis (rorqual), vlag (flag), vleugel (wing), voetbal (soccer), vonk (sparkle), vreugde (joy), vriezer (freezer), vuist (fist), vulkaan (volcano), wasbak (washbasin), worm (worm), zeep (soap), zon (sun).

L. Picture Names of chapter 5 with English translations in parenthesis

aambeeld (anvil), aardbei (strawberry), accordeon (accordion), ajuin (onion), artisjok (artichoke), barbecue (barbecue), bever (beaver), bokaal (jar), boormachine (drill), brievenbus (mailbox), broodrooster (toaster), cactus (cactus), denneappel (pinecone), deurknop (doorknob), dinosaurus (dinosaur), dolfijn (dolphin), eikel (acorn), eland (moose), enveloppe (envelope), fornuis (stove), fototoestel (camera), gewei (antlers), glijbaan (slide), hagedis (lizard), helicopter (helicopter), hengel (fishingpole), hoefijzer (horseshoe), jojo (yoyo), kangoeroe (kangaroo), kapstok (hanger), katapult (slingshot), kever (bug), kinderwagen (stroller), kleerkast (closet), koning (king), kruiwagen (wheelbarrow), kurketrekker (corkscrew), lippen (lips), magneet (magnet), mixer (mixer), olifant (elephant), paddestoel (mushroom), palmboom (palmtree), paperclip (clip), parachute (parachute), pelikaan (pelican), penseel (paintbrush), pleister (bandaid), racket (rocket), robot (robot), rolschaats (rollerskate), scheermes (razor), selder (celery), sneeuwman (snowman), spaarvarken (piggybank), spinneweb (spiderweb), statief (tripod), strijkijzer (iron), struisvogel (ostrich), tandenborstel (toothbrush), tennisracket (tennisracket), tractor (tractor), trechter (funnel), trommel (drum), tuinslang (hose), ventilator (fan), verrekijker (telescope), vingerhoed (thimble), vlieger (kite), vliegtuig (airplane), vlinder (butterfly), vos (fox), vrachtwagen (truck), wandelstok (cane), wasknijper (clothespin), weegschaal (scale1), wereldbol (globe), zaklamp (flashlight), zeehond (seal), zeilboot (sailboat).

M. Lures of the recognition memory test in chapter 5 with English translations in parenthesis

prikbord (notice board), servet (napkin), slagboom (barrier), anijs (anise), hangar (airdock), kous (stocking), bladzij (page), lantaarnpaal (street latern), vis (fish), hek (fence), militair (soldier), wolk (cloud), muts (beanie), veulen (foal), hoepel (hoop), trap (stairs), vork (fork), brandstapel (pyre), God (God), woningen (houses), bazaar (bazaar), tenen (toes), tabak (tobacco), pruik (wig), discus (disc), moord (murder), blad (leaf), wiel (wheel), blocnote (notebook), borrel (drink), potlood (pencil), wang (cheek), kado (present), hamer (hammer), koe (cow), maan (moon), tafel (table), papegaai (parrot), frisbee (frisbee), iglo (iglo), emmer (bucket), banaan (banana), speer (javelin spear), ballon (balloon), tomaat (tomato), vinger (finger), golfclub (golf club), chef (chef), bus (bus), klep (rattle), schakelaar (switch), stokbrood (french bread), kraal (bead), haak (hook), neef (cousin nephew), sprei (bed cover), poort (gate), appel (apple), plectrum (plectrum), kurk (cork), parel (pearl), haas (hare), sesam (sesame), stuur (steering wheel), ding (thing), roos (rose), pijp (pipe), boerderij (farm), kip (chicken), drogisterij (drugstore), portemonnee (wallet), bord (plate), sabel (sabre), koekje (biscuit), balkon (balcony), pantserwagen (armoured car), lamp (lamp), ananas (pineapple), web (web), pipet (pipette), bloemperk (flowerbed), insekt (insect), trompet (trumpet), LP (record), pompoen (pumpkin), prei (leek), vulkaan (vulcano), schommelstoel (rocking chair), zon (sun), mes (knife), lucifer (match), bloem (flower), kruk (stool), ijsje (icecream), cocon (cocoon), brief (letter), sloot (ditch), parkeerplaats (parking lot), deur (door), flessen (bottles), tulpen (tulips), longen (lungs), lepel (spoon), schaar (scissors), vriezer (freezer), brug (bridge), spoor (track), deeg (dough), armband (bracelet), ladder (ladder), bestek (cutlery), hoed (hat), aapje (monkey), leuning (banister), mie (noodles), boot (boat), kennis (aquaintance), caravan (caravan), wasbak (washbasin), file (traffic jam), sofa (sofa), kegel (cone), doorn (thorn), puzzel (puzzle), beugel (braces), lezer (reader), eindstand (final score), duim (thumb), hut (hut), fluit (flute), studente (female student), brandhout (fire wood), flap (flap), douche (shower), ring (ring), worm (worm), kist (chest), molen (mill), inktstel (inkstand), vonk (sparkle), fakkel (torch), sjaal (scarve), peddel (paddle), blok (block), been (leg), eend (duck), knie (knee), hersenen (brains), pet (cap), ruitenwisser (windshield wiper), sinaasappel (orange), veld (field), boetiek (clothes boutique), muisje (mouse), gerecht (dish), chocoladereep (chocolate bar), gids (guide), batterij (battery), pit (pit), banjo (banjo).

Nederlandse Samenvatting

Het voeren van gesprekken is een belangrijk maar ingewikkeld onderdeel van ons leven. De meeste mensen hebben geen enkele moeite met converseren en ervaren dit als eenvoudig en plezierig. Tijdens een gesprek wissel je tussen de rol van een luisteraar en dat van een spreker. Deze wisseling van luisteren naar spreken en terug naar luisteren wordt het overnemen van een spreekbeurt (*turn-taking*) genoemd. Op basis van verzamelingen van opgenomen gesprekken hebben onderzoekers geconcludeerd dat *turn-taking* een zeer snel proces is. Ze berekenden dat het een spreker meestal slechts 300 milliseconden (ms) kost om de beurt over te nemen en een antwoord te geven. Dit is een verrassende conclusie, omdat psycholinguïsten, die de productie van spraak op een experimentele manier onderzoeken, hebben laten zien dat de productie van een enkel woord (de naam van een plaatje) al 600 milliseconden (ms) in beslag kan nemen. Voor het uitspreken van een zinsdeel is soms zelfs meer dan een seconde nodig. Hoe is het mogelijk dat sprekers zo snel hun spreekbeurt kunnen overnemen (rond de 300 ms), terwijl het voorbereiden van de reactie zelf veel meer tijd kost (600 ms of meer)? Dit is één van de kernvragen van de de psycholinguïstiek.

Eén manier om zo snel te kunnen reageren is om je reactie al voor te bereiden terwijl je nog aan het luisteren bent. Op die manier heb je alvast een gedeelte van de reactie voorbereid tegen de tijd dat je aan de beurt bent om te spreken. Dit idee werd onderzocht in Hoofdstuk 2. Ook werd in dit hoofdstuk onderzocht of de moeilijkheid van de reactie die je voorbereid een rol speelt in hoe snel je kunt reageren. In dit proefschrift wordt de term ‘efficiëntie van het plannen’ gebruikt om deze manipulatie te beschrijven. In Experiment 1 hoorden de proefpersonen één van de volgende zinnen: '*En het plaatje rechts/links, kan je beschrijven wat daar gebeurt?*' (de vroege timing conditie), of: '*En kan je beschrijven wat er gebeurt op het plaatje rechts/links?*' (de late timing conditie). Als proefpersonen de eerste zin hoorden, konden ze al vroeg hun reactie plannen, namelijk meteen nadat ze het woord 'rechts/links' hoorden. Als proefpersonen de tweede zin hoorden, konden zij pas laat hun reactie plannen, omdat het woord 'rechts/links' pas aan het einde van de zin kwam. De proefpersonen moesten reageren door één van de twee plaatjes (links of rechts) op het computerscherm te beschrijven. Bijvoorbeeld: *'Het varken gaat naar de hamburger en de theepot'*. Om de voorbereiding moeilijker te maken, verschenen de plaatjes op het scherm soms rechtop, en soms op de kop. Experiment 1 liet zien dat proefpersonen sneller reageerden wanneer ze vroeg konden beginnen met plannen (dus bij de zinnen '*En het plaatje rechts/links, kan je beschrijven wat daar gebeurt?*'), vergeleken met zinnen waarbij ze pas aan het einde van de zin met de planning konden beginnen (dus bij de zinnen '*En kan je beschrijven wat er gebeurt op het plaatje rechts/links?*'). Dit betekent dat, wanneer de situatie het toelaat, sprekers inderdaad alvast hun reactie plannen terwijl ze nog aan het luisteren zijn. Deze strategie zorgt ervoor dat ze sneller de spreekbeurt kunnen overnemen. Echter, wanneer de planning wordt bemoeilijkt door de plaatjes op de kop te presenteren, dan lijken de proefpersonen hier niet door te worden afgeremd.

Experiment 2 was identiek aan Experiment 1, behalve dat de voorwerpen in de plaatjes nu zo gepresenteerd werden dat ze beschreven moesten worden door een complex zinsdeel, gevolgd door een simpel zinsdeel. Bijvoorbeeld: '*Het varken en de hamburger (complex zinsdeel) gaan naar de theepot (simpel zinsdeel)'* in tegenstelling tot *'Het varken (simpel zinsdeel) gaat naar de hamburger en de theepot (complex zinsdeel)'* uit Experiment 1. Net als in Experiment 1 waren proefpersonen sneller met reageren wanneer ze 'rechts/links' vroeg hoorden, vergeleken met wanneer het laat in de zin kwam, wat suggereert dat, wanneer de situatie het toelaat, sprekers hun reactie al plannen terwijl ze nog aan het luisteren zijn. Echter, in tegenstelling tot Experiment 1, had het manipuleren van de moeilijkheid voor het plannen van de reactie (door middel van het op de kop presenteren van de plaatjes) nu wel een effect op de snelheid waarmee proefpersonen reageerden. Er was een duidelijke invloed van efficiëntie van planning: in de late timing condities (waar proefpersonen moesten reageren op de vraag question '*En kan je beschrijven wat er gebeurt op het plaatje <u>rechts/links</u>?*') duurde het langer voordat ze reageerden wanneer de plaatjes op de kop waren gepresenteerd, dan wanneer ze rechtop stonden. Dit suggereert het volgende: In Experiment 1 konden proefpersonen al snel het eerste simpele zinsdeel voorbereiden (zoals '*Het varken*') ongeacht de oriëntatie van de plaatjes, en stelden zij waarschijnlijk de planning van het complexere zinsdeel ('*de hamburger en de theepot'*) uit tot na de start van hun reactie. In Experiment 2 was een dergelijke strategie onmogelijk aangezien de reactie als eerste het complexe zinsdeel ('*<u>Het varken en de hamburger</u>')* vereiste. De resultaten van Hoofdstuk 2 tezamen suggereren dat de tijd nodig voor het plannen van spraak een directe invloed uitoefent op de timing van het *turn-taken*. De efficiëntie van het plannen (hier onderzocht door middel van de hoeveelheid spraak dat voorbereid moest worden) speelt mogelijk een rol afhankelijk van of er vroeg of laat gepland wordt. Dus, wanneer sprekers geen tijd hebben om van tevoren hun reactie te plannen tijdens het luisteren, dan heeft de hoeveelheid spraak een effect op de timing van hun *turn-taking*.

De resultaten van Hoofdstuk 2 suggereren dat sprekers hun reactie voorbereiden tijdens het luisteren als dit mogelijk is. Deze conclusie werpt echter de volgende vraag op: Als sprekers beginnen met het voorbereiden van hun reactie tijdens het luisteren, dan zijn zij twee taken tegelijkertijd aan het uitvoeren: luisteren en spraak voorbereiden. Onderzoekers omschrijven deze noodzaak om aan twee taken tegelijkertijd aandacht te schenken als *dual tasking*. Het is bekend dat wanneer mensen twee taken tegelijkertijd uitvoeren, dit ten koste gaat van één van beide taken. Betekent het *dual tasken* tussen luisteren en spraak voorbereiden dat er nadelige gevolgen zijn voor één of beide processen? Of, in andere woorden: kunnen we nog wel aandachtig luisteren als we tegelijkertijd bezig zijn met het plannen van onze reactie? En omgekeerd?

In Hoofdstuk 3 werd er onderzocht of het plannen van spraak tijdens luisteren ervoor zorgt dat het luisteren minder goed gaat. Twee experimenten evalueerden luisterkwaliteit tijdens het plannen van spraak. In Experiment 1 en Experiment 2 luisterden de proefpersonen naar losse woorden terwijl ze een woord planden (in de plan-taak) of niet (in de niet-plan-taak). Om het plannen van woorden uit te lokken, verschenen plaatjes die later beschreven moesten worden op het beeldscherm tijdens het luisteren. In de niet-plan-taak verscheen er een lijntekening zonder betekenis, die dus ook niet beschreven kón worden. Hierdoor was de hoeveelheid visuele informatie gelijk in beide taken. Na afloop van de taak deden de proefpersonen een herkennings-geheugentaak: ze hoorden een lijst met woorden en moesten aangeven of ze het woord hadden gehoord tijdens vorige taak. Het verschil tussen de twee experimenten zat in het feit dat in Experiment 1 de proefpersonen van tevoren niet waren gewaarschuwd voor de herkenningstaak (de toevallige geheugenopslag modus), terwijl dit in Experiment 2 wel het geval was (de intentionele geheugenopslag modus). Tijdens beide experimentele fases werd de grootte van de oogpupillen van de proefpersonen gemeten. Eerder onderzoek heeft uitgewezen dat het uitvoeren van een lastige cognitieve taak af te meten is aan de grootte van de pupillen. Hoe meer inspanning geleverd wordt, hoe groter de verwijding van de pupillen.

De resultaten van beide experimenten laten zien dat de proefpersonen beter waren in het herkennen van woorden die ze eerder hadden gehoord in de niet-plan-taak, dan in de plan-taak. Bovendien waren de proefpersonen in Experiment 1 ook sneller in het juist herkennen van woorden in de niet-plan-taak, dan in de plan-taak. De toegenomen inspanning om de spraak te plannen was af te lezen aan de pupilwijdtes, met een grotere verwijding wanneer er spraak gepland werd tijdens het luisteren, dan wanneer de proefpersonen slechts hoefden te luisteren. Deze bevindingen laten zien dat het plannen van spraak tijdens het luisteren een cognitieve inspanning vergt die zorgt voor een afname in luisterkwaliteit.

De impact die het plannen van spraak tijdens luisteren heeft op de luisterkwaliteit kon niet ongedaan gemaakt worden, zelfs niet wanneer proefpersonen bewust de gehoorde woorden probeerden te herinneren (Experiment 2). Het is interessant dat het verplaatsen van de aandacht naar het luisteren de algehele prestatie tijdens de geheugentaak verbeterde , maar dat de proefpersonen tegelijkertijd meer versprekingen produceerden (in Experiment 2, vergeleken met Experiment 1). De resultaten van Hoofdstuk 2 laten duidelijk zien dat het plannen van spraak een negatief effect heeft op het verwerken van gehoorde woorden. Bovendien laten ze zien dat luisteraars de mogelijkheid hebben om prestaties op één taak strategisch te verbeteren, ten nadele van een andere taak.

Aangezien onderzoek naar gespreksverzamelingen heeft laten zien dat de meerderheid van de pauzes tussen spreekbeurten zeer kort is, lijkt het erop dat sprekers regelmatig worden blootgesteld aan het type verstoring tussen luisteren en spraak plannen zoals aangetoond in Hoofdstuk 3. Echter, alledaagse gesprekken bestaan bijna nooit uit alleen maar losse woorden. Dit roept de vraag op of het luisteren naar zinnen deze wisselwerking tussen luisteren en het plannen van spraak misschien verandert. De context van een zin zou deze invloeden van spraakplanning op luisterkwaliteit misschien gedeeltelijk kunnen helpen voorkomen. Ook zou deze zinscontext sprekers kunnen helpen met het voorspellen van gedeeltes van de gehoorde spraak, waardoor het plannen van spraak misschien minder inspanning kost.

Net als in Hoofdstuk 3, werd in Experiment 1 van Hoofdstuk 4 onderzocht hoe goed luisteraars kunnen omgaan met de dubbele last van het plannen van spraak tijdens het luisteren. In dit hoofdstuk luisterden de proefpersonen naar complete zinnen, in plaats van losse woorden. Om te testen of voorspelbaarheid de cognitieve last van het plannen van spraak tijdens luisteren misschien zou kunnen verlichten, werden de laatste woorden van de zinnen in óf een beperkende zinscontext óf een niet-beperkende zinscontext ingebed. Bijvoorbeeld, het woord ‘anker’ aan het einde van de zin werd eenmaal in de beperkende zinscontext ‘Het schip werd stevig vastgelegd met een anker’ ingebed, en eenmaal in de niet-beperkende zinscontext ‘Dat zware metalen stuk daar is een anker’. De proefpersonen planden woorden (in de plan-taak) of niet (in de niet-plan-taak), terwijl ze naar zinnen luisterden. Hierna voerden ze een herkennings-geheugentaak uit. De resultaten van Experiment 1 repliceerden de resultaten uit Hoofdstuk 3, maar nu in een zinscontext. Dus, de woorden aan het einde van de zin die gehoord werden tijdens het plannen van spraak werden slechter en langzamer herkend, vergeleken met de woorden die gehoord werden zonder dat er spraak gepland werd. Echter die impact van spraakplanning op de luisterkwaliteit werd niet verlicht wanneer de woorden aan het einde van de zin voorspelbaar waren. Dus, voorspelbare woorden aan het einde van de zin werden niet vaker herkend dan onvoorspelbare woorden, wanneer er werd gepland tijdens het luisteren. Aan de andere kant was de voorspelbaarheid wel van invloed op de snelheid waarin de proefpersonen de plaatjes benoemden. Net als in Hoofstuk 3 kon het plannen van spraak tijdens luisteren gelinkt worden aan een toegenomen cognitieve inspanning, aangezien de oogpupillen wijder waren wanneer er spraak gepland werd tijdens het luisteren, dan wanneer er alleen geluisterd werd.

Ook al komen de resultaten van Experiment 1 overeen met die van Hoofdstuk 3, de verwerkingslast die beide keren werd gevonden voor luisterkwaliteit hoeft niet persé veroorzaakt te zijn door het plannen van spraak, maar werd wellicht veroorzaakt door processen op een lager cognitief niveau, zoals object-herkenning. Om deze mogelijkheid te onderzoeken werd er een controle experiment uitgevoerd. Experiment 2 van Hoofdstuk 4 was identiek aan Experiment 1, behalve dat de proefpersonen nu de instructie kregen om naar de zinnen te luisteren en alleen passief de plaatjes te bekijken (de plaatjes hoefden in Experiment 2 dus *niet* beschreven te worden). In tegenstelling tot Experiment 1 was de pupilwijdte niet groter wanneer de plaatjes bekeken werden tijdens het luisteren (onthoud dat de plaatjes in dit experiment nooit benoemd hoefden te worden), dan wanneer hetzelfde onbeschrijfbare plaatje uit de niet-plan conditie van Experiment 1 bekeken werd. Deze uitkomst versterkt het beeld dat de gemeten pupilwijdten in Experiment 1 en in Hoofdstuk 3 toe te schrijven zijn aan de noodzaak om spraak te plannen tijdens het luisteren. Belangrijk om te vermelden is dat er geen verschil was in hoe goed de proefpersonen zich de woorden aan het eind van de zin herinnerden tijdens het bekijken van de plaatjes en tijdens het bekijken van het onbeschrijfbare plaatje. Dit laat zien dat het nadelige effect op de geheugenprestatie dat we observeerden in Experiment 1 en in Hoofdstuk 3 te wijten is aan het plannen van spraak tijdens het luisteren. De geobserveerde effecten in Hoofdstuk 3 en 4 (Experiment 1) kunnen dus inderdaad gelinkt worden aan planningsprocessen en niet aan object-herkenningsprocessen van een lager cognitief niveau.

Hoewel Hoofdstuk 3 en 4 lieten zien dat het plannen van spraak tijdens luisteren nadelige effecten kan hebben, waren deze conclusies gebaseerd op een *offline* (achteraf gemeten) maat van luisterkwaliteit, namelijk herinneringsgeheugen. De experimenten maten de impact van het plannen van spraak tijdens het luisteren door te kijken hoe goed proefpersonen woorden herinnerden die ze eerder hadden gehoord. Kan de impact van spraakplanning op luisterkwaliteit echter ook *online* gemeten worden, dus terwijl het gebeurt? Kunnen we *live* zien hoe de luisterkwaliteit verslechtert wanneer er tegelijkertijd spraak gepland moet worden?

In Hoofdstuk 5 werd er naar bewijs gezocht dat de impact van het gelijktijdig plannen van spraak op de luisterkwaliteit gemeten kan worden terwíjl het gebeurt. De luisterprestatie werd *online* gemonitord door middel van elektro-encefalografie (EEG), met de nadruk op het N400 component. Elektro-encefalografie is een niet-invasieve manier om de elektrische activiteit van de hersenen te meten. Deze activiteit geeft inzicht in de cognitieve processen die zich voltrekken tijdens het uitvoeren van de taak. Het N400 component wordt beschouwd als de elektrofysiologische handtekening van de verwerking van (semantische) taalbetekenis en kan worden opgeroepen door elke stimulus die mogelijk betekenis in zich draagt, zoals geschreven of gehoorde woorden, gebarentaal, plaatjes, omgevingsgeluid, enzovoort. Het is belangrijk om op te merken dat wanneer een stimulus wordt ondersteund door de context, de amplitude (het verschil tussen de hoogste en laagste waarde) van N400 wordt beperkt, een maat voor contextuele vergemakkelijking. Onderzoek heeft eveneens uitgewezen dat de amplitude van de N400 hoger wordt wanneer de stimulus, zoals bijvoorbeeld een woord aan het einde van een zin, op basis van de context minder verwacht, of zelfs gewoon incorrect is. Dit fenomeen wordt het N400 semantische anomalie-effect genoemd. We vroegen ons af of de auditieve invoer nog steeds op een grondig semantisch niveau verwerkt wordt (en dus een N400 oplevert), wanneer het woord aan het einde van de zin niet meer de volledige aandacht kan opeisen, zoals het geval is wanneer er op dat moment spraak gepland wordt.

Net als in Hoofdstuk 3 en 4 waren de proefpersonen bezig een woord te plannen (in de plan-taak) of niet (in de niet-plan-taak). Na afloop van de taak voerden zij een herkennings-geheugentaak uit. Net als in Hoofdstuk 4 keken we naar de mogelijke rol van voorspelbaarheid in het vergemakkelijken van de gelijktijdige processen van het plannen tijdens luisteren. Echter, in plaats van voorspelbare zinnen met onvoorspelbare zinnen te vergelijken, bevatten de gehoorde zinnen nu een voorspelbaar woord ('bij de thee krijgen we altijd een koekje') of een afwijkend woord ('bij de thee krijgen we altijd een muis') aan het eind. Dit soort zinnen worden vaak gebruikt in experimenten waar het N400 semantische anomalie-effect wordt onderzocht. Het gebruik van verwachte en afwijkende woorden aan het eind van de zinnen diende in dit experiment een dubbel doel: het stelde ons in staat om te zien hoe het N400 semantische anomalie-effect werd beïnvloed door het gelijktijdig plannen en luisteren, terwijl we tegelijkertijd konden testen of het horen van voorspelbare/verwachte inhoud de luisterkwaliteit kan helpen, ondanks dat er op dat moment gepland moest worden.

Event-Related-Potentials (*ERP*s, veranderingen in elektrische hersenactiviteit als gevolg van een gebeurtenis) die ontstaan tijdens het luisteren naar de woorden aan het einde van de zinnen lieten een substantiële vermindering van de amplitude van het N400 semantische anomalie-effect zien voor de woorden die gehoord werden tijdens het plannen van spraak, in vergelijking met woorden die gehoord werden terwijl er niet gepland hoefde te worden. Dit resultaat laat zien dat tijdens het plannen van spraak de verwerking van gehoorde spraak verminderd is, waardoor de semantische betekenissen niet volledig worden geactiveerd. Net als in Hoofdstuk 3 en 4, zagen we dat het herkenningsgeheugen voor woorden aan het einde van de zin slechter was in de plan-taak, dan in de niet-plan-taak. In navolging van Hoofdstuk, 3 (Experiment 1) en Hoofdstuk 4, zagen we dat woorden in de niet-plan-taak significant sneller als bekend werden aangemerkt, dan in de plan-taak. In tegenstelling tot Hoofdstuk 4, zagen we in dit experiment wel een effect tussen verwachte en afwijkende woorden op de luisterkwaliteit: als we alleen naar de woorden kijken in de plan-taak, dan zien we dat de door de context verwachte woorden beter werden herinnerd dan de semantisch afwijkende woorden. Deze overeenstemming tussen context en woord was ook van invloed op hoe snel de proefpersonen het plaatje beschreven, want na het horen van voorspelbare woorden werden de plaatjes sneller benoemd dan na het horen van afwijkende woorden.

Conclusie

Wat zijn de belangrijkste conclusies van de experimenten in dit proefschrift? Ten eerste, het beginnen met plannen van spraak tijdens het luisteren kan zorgen voor een kortere pauze tussen twee spreekbeurten (Hoofdstuk 2). En aangezien deze pauzes tussen gesprekspartners over het algemeen zeer kort zijn, kunnen we stellen dat zij vaak al hun reactie aan het plannen zijn terwijl ze nog naar de ander luisteren. Deze bevinding is door een aantal andere recente studies bevestigd. Ten tweede, wanneer sprekers maar heel weinig tijd hebben om hun reactie voor te bereiden tijdens het luisteren, wordt het *turn-taken* vertraagd door de hoeveelheid spraak die voorbereid moet worden (Hoofdstuk 2, Experiment 2).

Ten derde, het plannen van spraak tijdens het luisteren brengt een verwerkingslast met zich mee. Sprekers herinneren zich de woorden die ze hoorden tijdens het plannen van spraak minder vaak dan woorden die ze hoorden terwijl ze geen spraak aan het plannen waren. Deze cognitieve last zagen we niet alleen *offline* in een herinnerings-geheugentaak (Hoofdstuk 3, 4 en 5), maar ook *online* (Hoofdstuk 5) in een EEG-experiment waar een significante afname in de amplitude van het N400 semantische anomalie-effect gevonden werd voor woorden gehoord tijdens het plannen van spraak vergeleken met woorden wanneer er niet gelijktijdig spraak gepland hoefde te worden.

Hoewel het niet duidelijk is welke luisterprocessen precies hinder ondervinden van het plannen van spraak, is het wel duidelijk dat er verwerkingscapaciteit bij het luisteren vandaan gehaald wordt en in plaats daarvan aan het plannen besteed wordt. Het feit dat de pupilverwijding groter was tijdens het plannen van spraak (dan wanneer er niet gepland werd) bevestigt verder dat het plannen van spraak de vraag naar cognitieve capaciteit vergrootte. Het is belangrijk om op te merken dat het controle-experiment in Hoofdstuk 4 bevestigde dat het gelijktijdig plannen van spraak (en dus niet het herkennen van een afbeelding) de oorzaak is van deze interferentie en bijbehorende cognitieve last.

Ten vierde, sprekers zijn in staat om hun algehele prestatie in de herinnerings-geheugentaak te verbeteren wanneer ze actief focussen op het luisteren, echter, zij onthouden woorden tijdens het plannen nog steeds slechter dan woorden wanneer ze niet hoeven te plannen. Tegelijkertijd beïnvloedde het actieve luisteren hun reacties: ze maakten hierdoor meer spreekfouten (Hoofdstuk 3).

Ten vijfde, voorspelbaarheid speelt een meervoudige rol tijdens het *dual tasken*. De experimentele manipulaties in dit proefschrift focusten op hoe voorspelbaarheid het luisteren zou kunnen vergemakkelijken, waardoor het minder vatbaar is voor interferentie van het plannen van spraak (Hoofdstuk 4 en 5). We vergeleken voorspelbare woorden aan het einde van een zin met onvoorspelbare (maar nog wel mogelijke) woorden aan het einde van een zin, maar vonden geen voordeel voor de luisterkwaliteit (Hoofdstuk 4), terwijl de vergelijking tussen verwachte en afwijkende woorden aan het einde van zin wel een voordeel opleverde voor het herinneren van verwachte woorden (Hoofdstuk 5). Beide soorten voorspelbaarheid maakten het proces in zijn geheel wel iets makkelijker, aangezien sprekers sneller reageerden wanneer ze een voorspelbaar of verwacht woord hoorden, in vergelijking met een onvoorspelbaar of afwijkend woord. Voorspelbaarheid kan in bepaalde omstandigheden dus het omgaan met de dubbele taak van luisteren en spraak plannen vergemakkelijken en de pauze tussen spreekbeurten verkorten.

Dit proefschrift heeft een gedeelte van de verwerkingsbeperkingen (vanwege de noodzaak om snelle en efficiënte planning van spraak te combineren met luisteren) die voorkomen tijdens het voeren van een gesprek blootgelegd. Om succesvol een gesprek te kunnen voeren, moeten sprekers in staat zijn om hun cognitieve capaciteit strategisch in te zetten tussen het plannen van hun eigen spraak en het luisteren naar hun gesprekspartner. De experimenten in de proefschrift leken weliswaar niet op een werkelijk gesprek, toch stelde deze strenge experimentele controle ons in staat om een aantal factoren aan te wijzen die een rol spelen bij het strategische inzetten van capaciteit door gesprekspartners. Om efficiënt een gesprek te kunnen voeren moet een gesprekspartner beslissen of hij/zij wil focussen op luisteren, met als gevolg dat reageren misschien langzamer gaat, of dat het geven van een snelle en precies-geformuleerde reactie belangrijker is. Een gesprekspartner kan dan ook gezien worden als een *multi-tasker* die zo efficiënt mogelijk zijn/haar cognitieve capaciteit moet verdelen (afhankelijk van vraag naar capaciteit en de doelen die de spreker voor ogen heeft) voor een optimaal gesprek.

Acknowledgements

Long ago, when I started my PhD at the Max Planck Institute for Psycholinguistics, I did not know that I had been assigned a quest in an -at most- unknown territory. As with many things in life, I would find out in the process … Tackling new challenges and learning new things had always given me pleasure. So, I took a deep breath and dived into the ocean to find out -frankly- more things than I had anticipated. I can now confidently say that every PhD track is not only a unique track of scientific learning and scientific development. It is also a track of personal(ity) development. That is, not only have I acquired fascinating and useful new insights on looking at research data; I have also discovered that when the situation demands it, the greatest strength can come to oneself. A strength also nourished by people around you that support you.

It has been a tough many years, but also inspiring and magical. During these years both my sons were born; and I cannot help but think that there is nothing I care more about, than them.

So ... finally!!! How many years? How many tears? How many laughs? Almost everyone can find the answer to the first question, but only a few to the second and third. And only some people can read through these acknowledgments and get the full meaning of my words. This thesis is dedicated to them. Because after all; we are all just people, and the length of the acknowledgments section will be analogous to my PhD-track duration!

First of all, I would like to thank my supervisor Peter Hagoort and co-supervisor Matthias Sjerps. Dear Peter I do not think I can put into words how grateful I am. Thank you! Thank you for your support, when it was most needed. It is not too much to say that without your key support none of this would have ended happily ever after. I am grateful for your trust in me and this work.

Matthias, this has been such a long journey and there have been quite some storms on the way. I am grateful for your support up to the end of this journey. You were always part of it, in person, or from far. I still remember your efforts to offer some feedback on experimental design and other things, while you were away in the USA. Thank you for introducing me into mixed effects modeling. Thank you for always providing prompt feedback to my manuscripts.

I would also like to thank my former promotor Antje Meyer. Thank you, Antje, for all the time and energy you spent on this project that started at a time when so little was known about the psycholinguistics of turn-taking in conversation.

Let me also thank the members of the reading committee Prof. James McQueen, Dr. Judith Holler and Dr. Roel Willems for evaluating this thesis.

Being part of the IMPRS made life more colourful and fun. Thank you, Rachel Sheer, Els den Os and Dirkje van der Aa for being there for us and creating a feeling of togetherness, as well as for organizing all these courses. Thank you, Kevin, for all the support when hope was almost lost and also for guiding me through the final steps of delivering my manuscript to the printer. I am also grateful to Mark Dingemanse for his support and advice as PhD Dean.

Thank you to the technical group for being there when needed. Special thanks to Alex Dukers, Ronald Fisher and Johan Weustink, for their assistance in setting up rather complex experiments. Alex, I think most if not all my e-mails to technical support for software etc were answered by you. Thank you! I also thank Tilman Harpe for drawing or editing quite some pictures for my speech production tasks.

Thank you also to the assistants who took off some of my workload by annotating endless recordings of speech production and even running a control study for me. Thank you, Annelies, for doing such a fine job in coordinating the assistants and for lending your voice to many of my stimuli. Thank you also for being the open and cheerful person you are. I always knew that I could stop by and talk with you if needed.

Thank you Joost for taking part in setting up the study of chapter 5, when Matthias was in the USA. Thank you also for being the kind-hearted person you are, because it helped remembering that people can be utmost successful, while still remaining grounded and kind. And it was also fun organizing the journal meetings for the Psychology of Language Group. Thank you also to my colleagues Agnieszka, Alastair, Aoju, Carmen, Chu, Conny, Dan, Esther, Evelyn, Falk, Florian, Francesca, Jana, Johanne, Lesya, Neil, Odette, Shiri, Suzanne and Zeshu.

To my current employer Marc I can only say thank you! Thank you, because in one of the most tackling moments in my life, when I had started to doubt my ability as a researcher, you offered me this postdoc position. This gave me not only the so much needed financial means to continue re-writing my manuscript, at the end of really long working and parenting days; it also gave me the psychological strength to keep going, despite all odds. Milestone by milestone. Decision by decision. I cannot thank you enough for this. Thank you also to Jasper, Seba, Snandan and other colleagues at the department for sharing my agony these last months.

Let me not forget to thank the notorious MPI volleyball team, which made sure that at least on Thursdays I left office in time! Thank you, Ad, Alex, Dirkje, Haiteng, Izabela, Kees, Reiner, Sarah, Thordis, and all other people whose names

inevitably don't easily come into my mind after so many years! I look forward to joining again at some point in the undefined future!

Dear Thordis (your name certainly comes into mind!), I was not sure whether to write in English or German. You are definitely one of those people who lived from close by all the stages I went through in this thesis. Especially the ones after Nestor was born, and then Orfeas. You are full of energy and calm at the same time. I don't know how you do that! Thank you for being there when having someone to talk mattered the most. I look forward to meeting again, once the covid-19 pandemic permits, and having our children play at home or at our favorite playground or during a concert in Forstgarten!

To my paranymphs! Dear Alma! Sharing the Party Office with you and Maartje, made life easier and much more fun! Thank you for always keeping an open ear, even when I was mumbling on my own, while doing some analysis or writing some document! I know this was very distracting for you, so thank you for your patience. You are a true friend, one of the most organized people I know and have an extremely high sense of responsibility. I am so happy that you will not move any further than Utrecht! This way we can always meet and spend some nice time together no matter what!

Dear Maartje! Thank you for agreeing on rather short notice to be my paranymph. You are one of the sweetest, most energetic (but also calm) and strong-willed people that I have ever met. Always active, always kind and always helpful to others. Remember the birthday lion card? I am sure you will have a memorable time in your new home! Can't wait to come visit with the kids!

While thanking my paranymphs I cannot, but think also of you dear Catalina! Kouklitsa! SARS-CoV-2 did not allow easy traveling, so you are my paranymph at heart! It's been so many years! So many memories! I still remember that day after the math class. You were such a great help! We 've always shared our fears and tears and this has been great support for me. You always reminded me that there is more to life than work. That caring about oneself is essential; and you never seized to amaze me, with all the things you could accomplish in your working and personal life. You are an excellent teacher. You are a giving, loving person, definitely far more capable than you sometimes seem to think!

In my years in the Netherlands (to be precise very nearby in little Kranenburg in Germany), I met many wonderful people, whom we became friends with. Some are still here and some have long left. Dear Adjmal, I/we am/are very lucky that you are one of those people that have NOT left Nijmegen! You are a great listener and have been one of the few people to visit us in the evenings and stay until the children are asleep, so that we can talk about our lives but also about the world, about science, about everything! Thank you for that. Thank you also for reminding me that a brilliant mind can come with a modest personality. Dear Vi! Another brilliant mind with a kind heart. When our ways first crossed, you offered me one of the best supervisions I have ever had for my master's thesis. Thank you for that, thank you for always being patient to explain and never taking things for granted. You also helped me to overcome a milestone for this thesis, by helping me finish the data analysis for chapter 5. You are always friendly, always warm-hearted and truly an inspiration. Especially after Nestor was born, it became rather hard to meet, but the few times we had the chance to socialize I always felt as if we had just talked yesterday. This is also the feeling I get whenever we talk online with Jonathan and Miyuki. You have been away for so many years, but it feels like yesterday, when we would come to your place and play all the fantastic board games that Jo would take the time to explain to us. Playing

board games and having delicious dinner cooked by you Miyuki, or the best pizza from you Jo, or BBQing! You people are such wonderful company, you are giving and caring and funny. Thank you for all the wonderful times that made me and Michalis feel that we belong somewhere, here in Nijmegen. Thank you for organizing the basketball games at the sports center of the Radboud and for making me participate for the first time in an improvisation group. It was so much fun! A big hug to you people! Dear Gabi, I want to thank you, for making me look at the funny side of things, whenever I would start complaining. I still hear you shout "Sveeetlaaaanaaa!!!" within the MPI, when we would meet at the corridors.

Ten years is an awfully long time to be separated from your friends and family in Greece. To make matters even worse, whilst I was abroad doing my master's and PhD the infamous crisis hit in Greece. I was left with my weird role as a remote watcher of dramatic personal stories evolving. There are no words to describe what this does to you and how hard it can be. Being away is hard enough already. Being away when your friends and family would need you the most, is beyond imagination. So here is to my friends in Greece! I love you! And I will switch to Greek to tell you what I want properly! Ελενίτσα μου! Φιλενάδα μου γλυκιά! Παίρνω μηδέν στο μάθημα "σηκώστε το τηλέφωνο και μιλήστε με την φίλη σας". Συγγνώμη για αυτό. Τόσα χρόνια μακριά κι όμως κάθε φορά που ερχόμαστε για διακοπές είναι σα να είμαστε συνέχεια μαζί. Σε αγαπώ και σε ευχαριστώ που είσαι στη ζωή μου. Από το δημοτικό! Αλεξάνδρα μου! Γλυκιά μου συμφοιτήτρια! Πόσα χρόνια; Πόσες κοινές εργασίες; Ακόμα σας θυμάμαι με τον Γιάννη να μας χαιρετάτε στο λιμάνι της Πάτρας. Περιμένω την στιγμή που θα έρθετε να μας παραλάβετε από το λιμάνι της Πάτρας! Δύσκολα χρόνια για εμάς έξω, αλλά πολύ δύσκολα και για σένα. Και όμως όπως πάντα έδειξες θάρρος, υπομονή και επιμονή. Σε αγαπώ και σε ευχαριστώ κουμπαρούλα μου που είσαι στη ζωή μου! Γιάννη, Βασίλη, Όμηρε, Κάριν, είστε υπέροχοι! Σανατανούλι! Με νιώθεις! Αν υπάρχει κάποιος δικός μου άνθρωπος που έχει παλέψει για να πετύχει αυτό που θέλει, είσαι εσύ! Θέληση και δύναμη ψυχής μέχρι εκεί που δεν πάει! Ένα υπέροχο πλάσμα που υπεραγαπώ! Και φρέσκια μανούλα με μια υπέροχη κορούλα και έναν υπέροχο Jörn. Ποιος το έλεγε ότι η χρονική αυτή στιγμή θα μας έβρισκε μόλις 1.5 ώρα απόσταση από σπίτι σε σπίτι;! Σε ευχαριστώ που είσαι στη ζωή μου! Γιάννα μου! Μπορεί από ταξίδι σε ταξίδι να σε ψάχνουμε, αλλά βασικά είσαι κι εσύ στην Ολλανδία. Νονά Γιάννα σε περιμένουμε πάντα με μεγάλη χαρά όταν καταφέρνεις να βρεις χρόνο να κοιμηθείς σπίτι μας! Αννούλα μου! Μας λείπετε πολύ. Μιχαέλα μου, ξέρεις καλά πώς είναι να αφήνεις τον τόπο σου για να σπουδάσεις. Και χαίρομαι πολύ που έχεις τώρα γυρίσει και τα πράγματα πάνε τόσο καλά! Είσαι ένας υπέροχος άνθρωπος με φοβερή αίσθηση του χιούμορ και σαρκασμό ακριβώς όσο πρέπει. Ανυπομονώ να τα πούμε πάλι από κοντά με την πρώτη ευκαιρία.

Liebe Elena, lieber Philip, lieber Alexander! Ich danke euch. Ihr seid einfach super liebe und hilfreiche Menschen, und es macht mir großen Spaß mit euch Zeit zu verbringen. Liebe Dagmar, Jornt und Nele, ich danke auch euch! Dagmar, du bist ein außerordentlich lieber Mensch, und ich freue mich jedes Mal, wenn wir es schaffen uns zu treffen!

Liebe Frau Starke, danke, dass wir über zehn Jahre in Ihrer wunderschönen Wohnung wohnen dürfen! Danke, dass Sie uns immer das Gefühl geben, dass wir dazu gehören, und dass wir, und besonders die Kinder, Ihren bildhübschen Garten genießen dürfen. Mit dem tollen Gemüsegarten, den Blumen, den Pferden, den Hirschen, den Hühnern, den Gänsen, und den Katzen! Ein tolles Erlebnis! Danke nochmal.

Acknowledgements

Αγαπημένοι μου συγγενείς, θείε Κώστα και θεία Γιούλη, θείε Λάμπρο και θεία Μαρίνα, θεία Πιπίτσα, γιαγιά Φρασίκλεια, σας ευχαριστώ από καρδιάς. Είναι υπέροχο να ξέρω, ότι κάθε φορά που θα είμαστε Ελλάδα, έχουμε την οικογένειά μας εκεί. Τις ρίζες μας! Μπορεί να λείπουμε 10+ χρόνια, αλλά κάθε φορά που συναντιόμαστε είναι σα να τα είπαμε μόλις χτες. Είστε υπέροχοι. Και μας λείπετε!

Αγαπημένη μου Κατερίνα, μανούλα! Πόσες φορές, πόσα χρόνια, ακούς για αυτό το διδακτορικό; Σε ευχαριστώ! Σε ευχαριστώ πρώτα απ’ όλα, γιατί χωρίς εσένα δε θα είχα τον Μιχάλη! Σε ευχαριστώ, γιατί μας έχεις στηρίξει με όλους τους τρόπους όλα αυτά τα χρόνια και είσαι δίπλα μας σε κάθε τρικυμία αλλά και χαρά. Είσαι ένας περήφανος, υπέροχος άνθρωπος, που δεν σταματά ποτέ να σκέφτεται πώς θα οργανώσει τα πάντα με τον καλύτερο τρόπο για όλους. Ξέρεις να είσαι διακριτική αλλά και δυναμική, όποτε το επιτάσσει η κατάσταση.

Mama und Papa, meine lieben Eltern, ich danke euch. Ich danke euch, dass ich mir immer aussuchen konnte, was ich machen will. Als ich mit 4 Jahren, wie verzaubert mir gewünscht habe, dass ich Ballett lerne, habt ihr mich hingeschickt, obwohl es teuer war. Die 14 Jahre haben sich gelohnt! Ihr stand auch hinter mir, als ich mit 6 Jahren zur Leichtathletik ging, mit 12 Jahren zur Sportschule wollte, als ... als ... Das war immer so. Es gab kein muss, wenn es um unsere Lebenswünsche ging. Und so kam es, dass ich nach meinem Kognitionswissenschaft Aufbaustudium in Griechenland, doch dann auswandern wollte. Da habt ihr wieder mit vollem Herzen geholfen. Nicht ein einziges Mal, all diese Jahre, habt ihr euch beschwert, dass ich hier bin und ihr in Griechenland. Und das waren keine einfachen Jahre. Auch nicht als Nestor und Orfeas geboren wurden. Ich weiss, dass es nicht einfach ist. Und ich weiss ganz klar, dass ich auch lieber bei euch sein will. Und dann noch diese Corona-Krise dazu … Danke Mamili, ευχαριστώ Βαγγέλη, μπαμπά! Ihr seid die Besten!

Acknowledgements

Παναγιώτη ξέρεις καλά από πέρα δώθε Ελλάδα-Γερμανία-Ελλάδα. Ξέρεις, όπως λέω και στην Ελένη, δεν είμαι καλή με τα τηλεφωνήματα. Αλλά είσαι ο μεγάλος αδερφός μου και σε αγαπώ. Σε ευχαριστώ που με έκανες πάντα να θυμάμαι ότι υπάρχει ζωή και έξω από το διδακτορικό! Το ίδιο και εσύ Σεμπάστιαν! Είσαι ο μικρός μεγάλος μου αδερφός! Όταν το να είμαστε κοντά στην υπόλοιπη οικογένεια ήταν αδύνατο, είχαμε τις υπέροχες επισκέψεις σας με τον Φίλιπ και η ζωή γινόταν πιο γλυκιά! Σας ευχαριστούμε, σας αγαπάμε και σας περιμένουμε ξανά!

Ζουζουνάκι μου! Αγάπη μου! Σειρά σου να μου πεις "άάάάντεεε ντεεεεε"! Τα λόγια δεν φτάνουν για να σου πω ευχαριστώ! 23 χρόνια έχουμε ζήσει μαζί τα πάντα. Η δική μου αναζήτηση για διδακτορικό μας έφερε στη Γερμανία και την Ολλανδία. Και εσύ ήσουν εκείνος που με έσπρωξε να κάνω το βήμα! Όπως σε τόσα πράγματα! Είσαι πάντα εκεί για μένα. Έχεις πάντα μια συμβουλή που θα οδηγήσει σε μια καλή λύση. Έτσι και όσες φορές χρειάστηκα την υποστήριξή σου στα χρόνια του διδακτορικού. Και δεν ήταν λίγες! Ειδικά σε εκείνη την κρίσιμη στιγμή δεν πιστεύω ότι θα τα κατάφερνα χωρίς εσένα. Είσαι ο τρυφερός βράχος μου. Ξέρω, ότι όταν τίποτα δεν είναι πια όπως θα το ήθελα, μπορώ να έρθω σε εσένα να το συζητήσουμε. Σε ευχαριστώ. Για μια υπέροχη ζωή μαζί, για 2 υπέροχα παιδάκια, για πολλά πολλά χρόνια ακόμα μαζί, όπως εμείς τα ονειρευόμαστε. Σειρά σου τώρα! Λίγος καιρός έμεινε και θα στηρίξεις και εσύ το διδακτορικό σου! Σε αγαπάω με όλο μου το είναι. Σε λατρεύω.

Mein kleiner grosser Nestor! Fünf Jahre alt bist du jetzt schon! Du weißt es noch nicht, aber du hast mir auch bei der Doktorarbeit geholfen. Die langen Tage bis Dezember 2014, als ich EEG Probanden untersuchte, warst du auch dabei! In meinem Bauch! Ich denke fast du kennst schon den Geruch des EEG Gels. Tja! Und dann bist du schon sehr früh zur Tagesmutter gegangen, damit Mama fertig schreiben kann. Das war nicht einfach. Ich danke dir mein kleiner großer! Mein Mäuschen, mein Knutschibär!

Mein kleiner Orfeas! Du bist jetzt schon 2 Jahre alt! Ich hatte alles anders geplant und gewollt aber das Leben ist wie zeichnen ohne Radiergummi. Das weißt du schon inzwischen. Du hast mir auch sehr geholfen. Hast ganz geduldig auf mich gewartet, als du mit 2 Monaten im Auto mit Papa auf mich warten musstest, dass ich von meinen Besprechungen zurückkomme. Du musstest auch sehr früh von einer Babysitterin betreut werden, damit Mama ihre Arbeit fertigkriegt. Ich danke dir mein kleiner! Mein Knutschibaby!

" 'Goodbye,' said the fox. 'And now here is my secret, a very simple secret: It is only with the heart that one can see rightly; what is essential is invisible to the eye' " (Antoine de Saint-Exupéry, The Little Prince). Indeed; if I had not put my heart at work for this, then I would have missed the point. Even in this scientific quest. Thank you all for making this happen!

Curriculum Vitae

Svetlana-Lito Gerakaki was born in 1979 in Athens, Greece. In 2002 she obtained her bachelor's degree in Germanistics, with specialization in linguistics, from the University of Athens. There she also obtained a master's degree (cum laude) in 2005 from the 2-year graduate program in basic and applied cognitive science. While completing these studies, she was involved as a junior researcher in a number of research projects on reading disability (Institute for Language and Speech Processing, Athens, Greece), theory of mind (Cognitive Science Laboratory, University of Athens), metacognitive strategies in reading comprehension (Department of Psychology, University of Athens) and analogical reasoning (ANALOGY project; what does it mean to be human (NEST program with FP6 of the EU).

Svetlana went on to obtain a second master's degree (2011) from the Cognitive Neuroscience Research Master of the Radboud University, in the Netherlands, with a specialization in psycholinguistics. In the same year she started her PhD research at the Max Planck Institute for Psycholinguistics, in Nijmegen, the Netherlands. She is now a postdoctoral researcher in the Auditory System, Multisensory Gaze Control and Executive Function Group (Donders Center for Neuroscience, Biophysics Department, Radboud University, Nijmegen, the Netherlands).

Publications

- Gerakaki, S., Sjerps, M.J., Rommers, J. Planning speech while listening, decreases the amplitude of the N400 semantic anomaly-effect. Manuscript in preparation.
- Gerakaki, S., Sjerps, M.J. Concurrent speech planning affects memory for single- and sentence-final-words. A role for predictability? Manuscript in preparation.
- Meyer, A. S., & Gerakaki, S. (2017). The art of conversation: Why it's harder than you might think. Contact Magazine. Retrieved from http://contact.teslontario.org/the-art-of-conversation-why-its-harder-than-you-might-think/.
- Roelofs, A., Dijkstra, T., & Gerakaki, S. (2013). Modeling of word translation: Activation flow from concepts to lexical items. *Bilingualism: Language and Cognition, 16*, 343 - 353. https://doi.org/10.1017/S1366728912000612
- Protopapas, A., Gerakaki, S. (2009). *Development of processing stress diacritics in reading Greek*. *Scientific Studies of Reading 13* (6): 453-483. https://doi.org/10.1080/10888430903034788
- Vosniadou, S., Skopeliti, I., Gerakakis, S., & Blitsas, P (2008). A*nalogies Can Facilitate the Understanding of Counter-Intuitive Expository Texts.* [Abstract] In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (p.1368). Austin, TX: Cognitive Science Society.
- Vosniadou, S., Skopeliti, I., & Gerakakis, S. (2007). *Understanding the role of analogies in restructuring processes*. In A. Schwering, U. Krumnack, K.U. Kuhnberger, & H. Gust (Eds.), *Analogies: Integrating Multiple Cognitive Abilities*, Publications of the Institute of Cognitive Science, *Volume 5-2007*, (39-43), Osnabruck, Germany.
- Protopapas, A., Gerakaki, S., & Alexandri, S. (2007). Sources of information for stress assignment in reading Greek. *Applied Psycholinguistics 28* (4): 695–720. https://doi.org/10.1017/S0142716407070373
- Protopapas, A., Gerakaki, S., & Alexandri, S. (2006). Lexical and default stress assignment in reading Greek. *Journal of Research in Reading 29* (4): 418–432. https://doi.org/10.1111/j.1467-9817.2006.00316.x

MPI Series in Psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turennout*
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences. *Andrea Weber*
17. Moving eyes and naming objects. *Femke van der Meulen*
18. Analogy in morphology: The selection of linking elements in Dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*

24. Contextual influences on spoken-word processing: An electrophysiological approach. *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure: Distinguishing contrastive from non-contrastive variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel B.M. Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch. *Mark Pluymaekers*
44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*
45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*

47. Crosslinguistic influence in first and second languages: Convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*
50. Lexical processing of morphologically complex words: An information-theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo, a Papuan language of the Solomon Islands. *Claudia Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*
53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*
56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*
57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*
58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*
59. Lexical interactions in non-native speech comprehension: Evidence from electro-encephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*
60. Processing casual speech in native and non-native language. *Annelie Tuinman*
61. *Split intransitivity in Rotokas, a Papuan language of Bougainville. Stuart Robinson*
62. Evidentiality and intersubjectivity in Yurakaré: An interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension: A neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*
65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*
66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*
67. The relevance of early word recognition: Insights from the infant brain. *Caroline Junge*

68. Adjusting to different speakers: Extrinsic normalization in vowel perception. *Matthias J. Sjerps*
69. Structuring language. Contributions to the neurocognition of syntax. *Katrien R. Segaert*
70. Infants' appreciation of others' mental states in prelinguistic communication: A second person approach to mindreading. *Birgit Knudsen*
71. Gaze behavior in face-to-face interaction. *Federico Rossano*
72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space**.** *Conny de Vos*
73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*
74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt Witteman*
75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*
76. Territories of knowledge in Japanese conversation. *Kaoru Hayano*
77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder*
78. Contributions of executive control to individual differences in word production. *Zeshu Shao*
79. Hearing speech and seeing speech: Perceptual adjustments in auditory-visual processing. *Patrick van der Zande*
80. High pitches and thick voices: The role of language in space-pitch associations. *Sarah Dolscheid*
81. Seeing what's next: Processing and anticipating language referring to objects. *Joost Rommers*
82. Mental representation and processing of reduced words in casual speech. *Iris Hanique*
83. The many ways listeners adapt to reductions in casual speech. *Katja Poellmann*
84. Contrasting opposite polarity in Germanic and Romance languages: Verum Focus and affirmative particles in native speakers and advanced L2 learners. *Giuseppina Turco*
85. Morphological processing in younger and older people: Evidence for flexible dual-route access. *Jana Reifegerste*
86. Semantic and syntactic constraints on the production of subject-verb agreement. *Alma Veenstra*
87. The acquisition of morphophonological alternations across languages. *Helen Buckler*
88. The evolutionary dynamics of motion event encoding. *Annemarie Verkerk*
89. Rediscovering a forgotten language. *Jiyoun Choi*

90. The road to native listening: Language-general perception, language-specific input. *Sho Tsuji*
91. Infants'understanding of communication as participants and observers. *Gudmundur Bjarki Thorgrímsson*
92. Information structure in Avatime. *Saskia van Putten*
93. Switch reference in Whitesands. *Jeremy Hammond*
94. Machine learning for gesture recognition from videos. *Binyam Gebrekidan Gebre*
95. Acquisition of spatial language by signing and speaking children: a comparison of Turkish sign language (TID) and Turkish. *Beyza Sümer*
96. An ear for pitch: on the effects of experience and aptitude in processing pitch in language and music. *Salomi Savvatia Asaridou*
97. lncrementality and Flexibility in Sentence Production. *Maartje van de Velde*
98. Social learning dynamics in chimpanzees: Reflections on (nonhuman) animal culture. *Edwin van Leeuwen*
99. The request system in Italian interaction. *Giovanni Rossi*
100. Timing turns in conversation: A temporal preparation account. *Lilla Magyari*
101. Assessing birth language memory in young adoptees. *Wencui Zhou*
102. A social and neurobiological approach to pointing in speech and gesture. *David Peeters*
103. Investigating the genetic basis of reading and language skills. *Alessandro Gialluisi*
104. Conversation Electrified: The Electrophysiology of Spoken Speech Act Recognition. *Rósa Signý Gisladottir*
105. Modelling Multimodal Language Processing. *Alastair Smith*
106. Predicting language in different contexts: The nature and limits of mechanisms in anticipatory language processing. *Florian Hintz*
107. Situational variation in non-native communication. *Huib Kouwenhoven*
108. Sustained attention in language production. *Suzanne Jongman*
109. Acoustic reduction in spoken-word processing: Distributional, syntactic, morphosyntatic, and orthographic effects. *Malte Viebahn*
110. Nativeness, dominance, and the flexibility of listening to spoken language. *Laurence Bruggeman*
111. Semantic specificity of perception verbs in Maniq. *Ewelina Wnuk*
112. On the identification of FOXP2 gene enhancers and their role in brain development. *Martin Becker*

113. Events in language and thought: The case of serial verb constructions in Avatime. *Rebecca Defina*
114. Deciphering common and rare genetic effects on reading ability. *Amaia Carrión Castillo*
115. Music and language comprehension in the brain. *Richard Kunert*
116. Comprehending Comprehension: Insights from neuronal oscillations on the neuronal basis of language**.** *Nietzsche H.L. Lam*
117. The biology of variation in anatomical brain asymmetries. *Tulio Guadalupe*
118. Language processing in a conversation context. *Lotte Schoot*
119. Achieving mutual understanding in Argentine Sign Language. *Elizabeth Manrique*
120. Talking Sense: the behavioural and neural correlates of sound symbolism. *Gwilym Lockwood*
121. Getting under your skin: The role of perspective and simulation of experience in narrative comprehension. *Franziska Hartung*
122. Sensorimotor experience in speech perception. *Will Schuerman*
123. Explorations of beta-band neural oscillations during language comprehension: Sentence processing and beyond. *Ashley Lewis*
124. Influences on the magnitude of syntactic priming. *Evelien Heyselaar*
125. Lapse organization in interaction. *Elliott Hoey*
126. The processing of reduced word pronunciation variants by natives and foreign language learners: Evidence from French casual speech. *Sophie Brand*
127. The neighbors will tell you what to expect: Effects of aging and predictability on language processing. *Cornelia Moers*
128. The role of voice and word order in incremental sentence processing. *Sebastian Sauppe*
129. Learning from the (un)expected: Age and individual differences in statistical learning and perceptual learning in speech. *Thordis Neger*
130. Mental representations of Dutch regular morphologically complex neologisms. *Laura de Vaan*
131. Speech production, perception, and input of simultaneous bilingual preschoolers: Evidence from voice onset time. *Antje Stoehr*
132. A holistic approach to understanding pre-history. *Vishnupriya Kolipakam*
133. Characterization of transcription factors in monogenic disorders of speech and language. *Sara Busquets Estruch*
134. Indirect request comprehension in different contexts. *Johanne Tromp*
135. Envisioning Language - An Exploration of Perceptual Processes in Language Comprehension. *Markus Ostarek*

136. Listening for the WHAT and the HOW: Older adults' processing of semantic and affective information in speech. *Juliane Kirsch*
137. Let the agents do the talking: on the influence of vocal tract anatomy on speech during ontogeny and glossogeny. *Rick Janssen*
138. Age and hearing loss effects on speech processing. *Xaver Koch*
139. Vocabulary knowledge and learning: Individual differences in adult native speakers. *Nina Mainz*
140. The face in face-to-face communication: Signals of understanding and non-understanding. *Paul Hömke*
141. Person reference and interaction in Umpila/Kuuku Ya'u narrative. *Clair Hill*
142. Beyond the language given: The neurobiological infrastructure for pragmatic inferencing. *Jana Bašnáková*
143. From Kawapanan to Shawi: Topics in language variation and change. *Luis Miguel Rojas-Berscia*
144. On the oscillatory dynamics underlying speech-gesture integration in clear and adverse listening conditions. *Linda Drijvers*
145. Understanding temporal overlap between production and comprehension. *Amie Fairs*
146. The role of exemplars in speech comprehension. *Annika Nijveld*
147. A network of interacting proteins disrupted in language-related disorders. *Elliot Sollis*
148. Fast speech can sound slow: Effects of contextual speech rate on word recognition. *Merel Maslowski*
149. Reason-giving in everyday activities. *Julija Baranova*
150. Speech planning in dialogue - Psycholinguistic studies of the timing of turn taking. *Mathias Barthel*
151. The role of neural feedback in language unification: How awareness affects combinatorial processing. *Valeria Mongelli*
152. Exploring social biases in language processing. *Sara Iacozza*
153. Vocal learning in the pale spear-nosed bat, Phyllostomus discolor. *Ella Lattenkamp*
154. The effect of language contact on speech and gesture: The case of Turkish-Dutch bilinguals in the Netherlands. *Elif Zeynep Azar*
155. Language and society: How social pressures shape grammatical structure. *Limor Raviv*
156. The moment in between: Planning speech while listening. *Svetlana-Lito Gerakaki*