Repetita iuvant?

Studies on the role of repetition priming as a supporting mechanism during conversation

Funding body
This research was conducted at the Max Planck Institute for Psycholinguistics (Nijmegen, The Netherlands) with support from the Max Planck Society for the Advancement of Science (www.mpg.de/en).

International Max Planck Research School (IMPRS) for Language Sciences
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

The MPI series in Psycholinguistics
Initiated in 1997, the MPI series in Psycholinguistics contains doctoral theses produced at the Max Planck Institute for Psycholinguistics. Since 2013, it includes theses produced by members of the IMPRS for Language Sciences. The current listing is available at www.mpi.nl/mpi-series

Cover design by Federica, background picture *bytriangle* on Pixabay
Printed and bound by Ipskamp Drukkers, Enschede

Repetita iuvant?

Studies on the role of repetition priming as a supporting mechanism during conversation

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 voor promoties

in het openbaar te verdedigen op

dinsdag 24 januari 2023

om 10.30 uur precies

door

Federica Bartolozzi

geboren op 13 juli 1993

te Prato (Italië)

Promotor:

Prof. dr. Antje S. Meyer

Copromotoren:

Dr. Suzanne R. Jongman (Tilburg University)

Dr. Ashley G. Lewis

Manuscriptcommissie:

Prof. dr. Herbert J. Schriefers

Prof. dr. Linda R. Wheeldon (Universitetet i Agder, Noorwegen)

Dr. Evelien S. Heijselaar

To my parents
Ai miei genitori

Contents

1 | General introduction

1.1 Introduction

Conversations can be hard: think about a job interview or a phone call with the plumber, whom you are desperately asking how to stop the leak under your sink. In the case of a job interview, you probably listened carefully to the questions that the interviewer was asking, at the same time trying very hard to come up with clever and exhaustive answers. Similarly, during a phone call with a plumber, the aim is to understand the plumber's instructions as fast as possible and at the same time ask precise clarification questions in case anything is unclear.

The examples above show that conversations require the combination of multiple processes. Indeed, speakers need to retrieve words from memory and combine them into meaningful utterances, while listeners need to segment the sound stream into individual words and reconstruct the meaning of the sentence. These processes occur in rapid succession, and sometimes overlap. Swift turn-taking is necessary for successful conversations. Extremely long gaps might reflect hesitation to answer or distraction. For instance, in the example about the job interview, an unusually long gap between the question of the interviewer and the answer of the applicant could be taken as a sign that the applicant is not knowledgeable about the topic of the question. In average conversations, gaps between turns are about 200ms long, way less than average latencies in lab settings (from 600ms to 1550ms, see Levinson (2016)), and recent studies have shown that interlocutors start preparing their utterance while the other person is still speaking (Bögels, Magyari, & Levinson, 2015; Corps, Crossley, Gambi, & Pickering, 2018; Levinson & Torreira, 2015). This means that participants need to continuously switch between comprehension and production, which I will refer to as linguistic dual-tasking. While doing two things at the same time is often possible, the end result is not always satisfactory: for instance, drivers are more distracted when they are also talking on the phone, and listening to the news on the radio is more effortful when one is simultaneously engaged in another task such as cooking or reading a book.

Conversations require the execution of complex processes in a timely - and sometimes overlapping - manner, as shown above. Yet, engaging in a conversation often seems easy and effortless. According to an influential model of dialogue, the interactive alignment model (Pickering & Garrod, 2004), conversations are usually smooth because interlocutors develop a common understanding of what is being discussed, a phenomenon called alignment. Going back to the example of the phone call with the plumber, understanding their instructions is much easier if the person receiving them already knows which specific valves or pipes the plumber is talking about. According to the interactive alignment model, alignment in conversation gradually increases due to priming, a well-known psychological phenomenon whereby the use of a specific part of an utterance (e.g., a word) increases the chance that the same part will be repeated.

This dissertation focused on two mains strands of research. On the one hand, I drew inspiration from linguistic dual-tasking studies to investigate possible interference between production and comprehension processes. On the other hand, I focused on priming, and tested its use and resilience in linguistic dual-tasking settings and natural conversation. The aim was to evaluate whether priming can indeed ease demands of comprehension and production and facilitate conversation. In the next paragraphs, I first report recent findings on the interference between production and comprehension during linguistic dual-tasking, and then detail the interactive alignment model. In the last paragraph, I focus on the main research questions of the following chapters and describe the structure of this dissertation.

1.2 Interference between production and comprehension

Carrying out two tasks at the same time is usually associated with performance decrements in one or both tasks (Boiteau, Malone, Peters, & Almor, 2014; Strayer & Johnston, 2001). For instance, it is well known that talking on the phone while driving can cause distraction and lead to car accidents, and we have all experienced situations in which speaking is more effortful because we are currently doing something else at the same time, for instance trying to figure out a new recipe. Recently, psycholinguistic studies have focused on investigating whether the combination of production and comprehension tasks is also a form of dual-tasking, and tested whether the overlap between these processes can yield performance decrements in production, comprehension, or both. In these studies, participants usually heard speech (e.g., questions, tones, syllables) while carrying out a production task (e.g., answering questions or naming pictures: Bögels, Casillas, and Levinson (2018); Bögels et al. (2015); Fargier and Laganaro (2016, 2019). Interference between production and comprehension is usually measured by comparing such experimental conditions to a condition in which participants perform the critical task in isolation (e.g., name pictures or listen to words without any concurrent tasks).

Studies that focused on the effect of simultaneous comprehension on production found that picture-naming is delayed when interlocutors are simultaneously listening to speech (e.g., Bögels et al., 2015; Fargier & Laganaro, 2016). In other words, speech planning is less efficient during linguistic dual-tasking than during single-tasking. Whether production processes can hinder comprehension has been less studied. Electrophysiological studies using syllables in the comprehension task and single words in the production task showed that early auditory ERP components elicited by the syllables can be delayed when the comprehension task is performed together with a production task rather than isolation (Daliri & Max, 2016; Fargier & Laganaro, 2019).

Concerning semantic processing, there is some evidence that comprehending a question while preparing a spoken response yields a reduced N400 (Bögels et al., 2018), an ERP traditionally associated with semantic processing (Kutas & Federmeier, 2011). More specifically, shorter production latencies in Bögels et al. (2018) correlated with a decreased amplitude of the N400, suggesting that starting to prepare an answer while the other person is still speaking is associated with less efficient comprehension. However, the reduction in the N400 amplitude as a consequence of linguistic dual-tasking was only evident in the correlation but no effect of linguistic dual-tasking on the mean amplitude of the N400 was found.

In sum, the evidence so far suggests that linguistic dual-tasking is associated with decrements in production, but effects on comprehension require further studies. In particular, it stills needs to be addressed whether and to what degree semantic processing of comprehended words is affected by a concurrent task. This question is addressed in *Chapter 3* and *Chapter 4* of this dissertation.

1.3 The interactive alignment model

While the section above highlighted that combining comprehension and production could result in processing costs, conversations in our daily lives appear swift and smooth. This suggests that the cognitive system does not simply engage in production and comprehension processes, but makes use of additional mechanisms to mitigate the negative consequences of linguistic dual-tasking. According to an influential model (Pickering & Garrod, 2004), this mechanism is interactive alignment. More specifically, Garrod and Pickering (2009) argue that, according to this model, conversations are effortless because *"communicators come to understand relevant aspects of the world in the same way as each other. In other words, they align their representation of the situation under discussion"*. In the original version of the model, alignment was achieved through priming, an *automatic* mechanism whereby interlocutors tend to repeat particular aspects of their own and each other's utterances (Pickering & Garrod, 2004). For instance, interlocutors are more likely to use the word *couch* rather than *sofa* after hearing the word *couch*. Priming can arise at various levels (e.g., lexical choice, phonology, syntax, word meaning, etc.) and priming at one level increases priming at other levels. For instance, repetition of a word boosts the repetition of syntactic structures, a phenomenon known as lexical boost (Hartsuiker, Bernolet, Schoonbaert, Speybroeck, & Vanderelst, 2008; Traxler, Tooley, & Pickering, 2014).

While the original version of the interactive alignment account focused on priming as the main mechanism for achieving alignment, in a subsequent version of the model (Garrod & Pickering, 2013a, 2013b), Pickering and Garrod suggested that alignment is tightly linked to prediction. During conversation interlocutors tend to covertly imitate and predict each other's utterances at different levels (e.g., lexical choice, phonology, syntax, etc.). Given that - according to the interactive alignment model - representations in production and comprehension are tightly linked, predicting upcoming words while another person is speaking will make it more likely for the upcoming speaker to repeat those words in a subsequent utterance, therefore leading to more alignment. In addition, more alignment also affords better predictions, because interlocutors have a similar understanding of what is being discussed. In the last few years, the relationship between prediction and priming has been object of investigation in theoretical and experimental studies (e.g., Rabovsky, Hansen, & McClelland, 2018). Lab experiments focused on whether the ability to predict certain words can impact on subsequent repetition priming effects (Hodapp & Rabovsky, 2021; Rommers & Federmeier, 2018b). The evidence suggests that unpredictable words are associated with stronger priming effects, while predictable words do not benefit much from repetition priming.

The interactive alignment model makes straightforward predictions concerning the role of priming in conversation, and makes claims about the relevance of prediction to foster alignment. Yet, in order to work as a supporting mechanism in conversation, priming must be strong enough to occur under the conditions prevailing in conversation. For instance, words in conversations can be repeated immediately but also after some delay. Furthermore, words are often embedded in sentences, which may affect their effectiveness as primes. Finally, as described in the previous paragraph, production and comprehension can overlap in conversation. Is repetition priming still effective when comprehension processes are made more complex and more effortful? This is the research question that I asked in *Chapter 2*, *Chapter 3*, and *Chapter 4*.

Chapter 5 is dedicated to investigating repetition priming during naturalistic comprehension. Furthermore, given the relevance that the interactive alignment account gives to both priming and prediction to achieve priming, *Chapter 5* explores the relationship between the two mechanisms. Earlier work (Hodapp & Rabovsky, 2021; Lai, Rommers, & Federmeier, 2021; Rabovsky et al., 2018; Rommers & Federmeier, 2018a, 2018b) suggests that the size of repetition priming effects may depend on how easily a word is predicted upon first presentation. In other words, repetition priming in conversation might work as an error-driven mechanism aimed at correcting mispredictions. If this is the case, the relationship between priming and prediction should be evident not only in lab settings but also in conversation and naturalistic comprehension. An attempt to test this relationship is presented in *Chapter 5*.

1.4 Research questions and structure of dissertation

As detailed in the previous paragraphs, this dissertation focused on two main strands of research. First of all, I investigated possible negative consequences of linguistic dual-tasking on comprehension processes, with a focus on online comprehension and resulting linguistic representations. The second strand of research concerned repetition priming, specifically, its role in supporting spoken interactions under conditions relevant for conversation.

The type of priming that was investigated in the next chapters is repetition priming, i.e., the repetition of words. Repetition priming has been object of study in the psychological literature since at least the early 1950s. Given that repetition priming has been studied in various domains, modalities and settings, it is a perfect study case to determine whether priming can indeed work as a supporting mechanism in conversation.

This dissertation is structured as follows:

Chapter 2 lays the foundations of the work carried out in the following chapters. In this chapter, I reviewed repetition priming studies that used production and comprehension tasks. The aim of the review was to determine whether repetition priming can, in principle, support conversation. In order to answer this question, I focused on three factors that might affect the occurrence and magnitude of priming, namely lag, embedding, and linguistic dual-tasking. *Lag* refers to the distance between prime and target. Determining whether and how lag - in time and intervening trials - impacts on repetition priming is relevant because in conversations words might be repeated immediately or after one or more turns. In order to work as an efficient supporting mechanism, priming should therefore occur even when words are not immediately repeated. *Embedding* refers to the sentence of which the prime words are part. While repetition priming studies often employ single words, conversations usually include longer utterances. It is therefore important to evaluate how repetition priming is affected when prime words are presented in sentences rather than in isolation. The concept of *linguistic dual-tasking* has been explained earlier in this chapter. Since comprehension and production can overlap, priming - as a supporting mechanism - should occur even when attention is split between these processes.

In *Chapter 3*, I described three experiments investigating whether repetition priming from comprehension to production is modulated by a concurrent production task. The answer to this research question is relevant for theories of speech planning in conversation because it enables us to evaluate whether linguistic dual-tasking can hinder the generation of linguistic representations obtained during comprehension or the access to these representations when words are repeated. In other words, in this chapter I tested whether repetition priming is resilient to linguistic dual-tasking. In all experiments, participants heard a sentence containing a prime word and then - immediately or after a few trials - named a target picture. The critical manipulation was whether or not the participants were preparing to name a distractor picture while hearing the prime word.

In *Chapter 3*, I used a behavioural paradigm to investigate interference from production to comprehension during linguistic dual-tasking. One drawback of this design is that it cannot provide information about the time course of comprehension processes. For this reason, *Chapter 4* presents an EEG study assessing more directly whether word processing is affected by concurrent production. In addition to focusing on linguistic representations and repetition priming, i.e., the end-result of comprehension, I also investigated how comprehension unfolds in real time when a production task is also being carried out. The chapter is divided in two main parts. In the first part, I investigated the effect of linguistic dual-tasking on online comprehension processes. Participants heard words in silence or during a picture-naming task while their EEG activity was being recorded. In the second part of the chapter I investigated whether - upon repetition - words initially presented under linguistic dual-tasking are processed differently from words initially presented in isolation.

While *Chapter 3* and *Chapter 4* investigated priming from production to comprehension, *Chapter 5* solely focused on priming in comprehension. The main research question was whether repetition priming can occur during naturalistic comprehension of stories and whether the embedding of the prime word has an impact on subsequent priming. To do so, I analysed a publicly available EEG dataset recorded while participants listened to an audio book. In the first part of the chapter, I replicated previous findings according to which predictability of a word affects the size of the N400 (Heilbron, Ehinger, Hagoort, & De Lange, 2019). Then I asked whether repetition of words is associated with a decrease in the N400, as widely shown in lab studies (Kutas & Federmeier, 2011). In the third part of the chapter, I asked whether repetition priming of a word is modulated by the predictability of the word at previous presentation.

In *Chapter 6* I summarised and discussed the results of previous chapters. Furthermore, I suggested avenues for further research and experiments that could not be designed and implemented due to the current COVID-19 pandemic.

2 | Constraints on repetition priming of spoken words: A systematic review

Abstract

Repetition priming is a widely studied phenomenon in psycholinguistics, which has also been proposed as an important supporting mechanism in conversation. To support conversation, priming needs to be effective when primes and targets follow each other immediately or with a delay, and when primes and targets appear in the same or different utterance contexts. In this review, we discuss the evidence pertaining to these constraints. In addition, priming should be effective when primes are processed under full as well as divided attention. We first describe the effect of intervening trials on priming, highlighting possible mechanisms for the decay of priming across lags. Then we discuss the effects of contextual embedding of the primes and explain how embedding primes in a sentence can affect repetition priming of both single word targets and targets embedded in sentences. In the concluding section, we discuss the findings of the review and propose ideas for further research.

Keywords: repetition priming; comprehension; production; conversation; lag; attention; context

2.1 Introduction

Repetition priming is a widely studied phenomenon in psycholinguistics, with modern studies spanning from the early 1950s to the present day (e.g., Barry, Hirsh, Johnston, & Williams, 2001; Barry, Johnston, & Wood, 2006; Forster & Davis, 1984; Heath et al., 2012a; Neisser, 1954; Wheeldon & Monsell, 1992; Wilder, Davies, & Embick, 2019). While early research on this phenomenon was driven by interest in dissociations between explicit and implicit memory, and in the mechanisms underlying lexical access (see Schacter (1987) for a review), later research also focused on repetition priming (but also semantic and, especially, syntactic priming) as a likely supporting mechanism in conversation (e.g., Garrod & Anderson, 1987; Garrod & Pickering, 2009; Pickering & Garrod, 2004, 2013).

Repetition priming, defined as the tendency to re-use or process faster previously presented words than new words, is a robust phenomenon and has been observed in production (e.g., Wheeldon & Monsell, 1992), comprehension (e.g., Wilder et al., 2019), and across modalities (e.g., Jongman & Meyer, 2017). Often, repetition priming is measured by having participants name pictures at study, at test, or both. Repetition priming for words can occur due to the facilitation of object identification processes, as in the case of faster recognition of a depicted object, or of word production processes (Francis, Corral, Jones, & Sáenz, 2008). In particular, the facilitation of word production processes arises at the interface between semantics and phonology (Monsell, Matthews, & Miller, 1992). The robustness and availability of priming across modalities make it a likely supporting strategy in conversation, whereby people may capitalise on their own and their interlocutor's utterances to facilitate and speed up production and comprehension.

Notwithstanding the huge amount of work on repetition priming and its purported role in everyday conversations, it is not clear to what extent priming can occur outside of lab settings. For instance, while naming a picture - a very common task in repetition priming experiments - simply requires identifying the portrayed object and retrieving its name from memory, planning a longer utterance requires the retrieval of multiple words, which need to be combined together and uttered at the right time (for instance, by waiting to start speaking while someone else is already doing it). It is an empirical issue whether - and to what extent - repetition priming occurs when these processes are performed almost simultaneously.

In this review, we describe the settings in which priming can and cannot occur, and we use this information to assess whether repetition priming can in principle work as a supporting mechanism in conversation. For this reason, we only reviewed studies that used spoken production and comprehension. The aim of this review is to determine whether priming can in principle work as a supporting mechanism in conversation. Finding that repetition priming can work in conversation-like settings would further validate models according to which repetition priming is used by speakers to reach common ground (i.e., common background knowledge; Garrod and Pickering (2009); Pickering and Garrod (2004, 2013)).

Language processing minimally includes three basic mechanisms: retrieving words from the mental lexicon, combining them into larger chunks, and using executive control to guide these processes (Chang, Dell, & Bock, 2006; Hagoort, 2005, 2013; Levelt, 1993; Roelofs, 2003). For instance, when comprehending an utterance, interlocutors must pay attention to the incoming input, retrieve relevant word forms from the mental lexicon and combine them into a meaningful sentence. To determine whether repetition priming can indeed be used to support conversation, it is therefore important to determine how priming effects can be influenced by these basic mechanisms.

In particular, we pursued three main questions. First, we investigated how priming is affected by a lag between prime and target. Investigations about how priming decays across lags can help us understand which mechanisms underlie priming and whether they are qualitatively different when the repetition occurs at short versus long lags. In addition, knowing how long priming can last can give us information about the extent to which this mechanism can be used in conversation, when words are repeated after one or more turns.

The second question concerns the effect of contextual embedding on repetition priming. Again, the interest in this question is two-fold since the answer enables us to determine 1) whether combining the prime with other words in a sentence hinders the subsequent retrieval of the prime as a single word or in a different context, and 2) whether the repetition of a word can facilitate word and sentence production when the sentential embedding changes across utterances, as is commonly the case in conversation.

The third question concerns the effect of attention on repetition priming. The answer to this question is important because it enables us to determine whether priming from comprehended to spoken words is an automatic process or whether it can be hindered by division of attention. As for the implications for conversation, research on turn-taking suggests that people may start to plan their response while their interlocutor is still speaking (Bögels et al., 2015; Corps et al., 2018; Fargier & Laganaro, 2016). This means that people might sometimes dual task between production and comprehension processes. By investigating attentional effects on priming, we can therefore determine whether priming also occurs when participants dual task between production and comprehension.

2.2 Methods

Different paradigms can be used to investigate repetition priming, both at study and at test (e.g., picture naming, fragment completion, perceptual identification tasks). In this review, we focused on repetition priming effects from spoken production or comprehension to production.

We searched relevant papers (articles and proceedings papers) using the Web of Science database, and using the following keywords: ("repetition priming" OR alignment) AND (lag OR delay OR context OR sentence OR text OR attention OR production OR comprehension), and only selecting fields of interest (psychology, psychology multidisciplinary, linguistics, psychology experimental, neurosciences, language linguistics). We repeated the search substituting the terms "repetition priming" and "alignment" with "common ground", again selecting the relevant fields (linguistics, language linguistics, psychology applied, psychology experimental, psychology multidisciplinary, psychology social, psychology, neurosciences). The first and second search yielded 2253 studies in total. We then excluded reviews, unpublished dissertations, studies about forms of priming other than word repetition (e.g., syntactic priming), clinical studies, studies about language development or ageing, and studies that employed reading tasks at both study and test. As for studies on the effect of context, we also excluded studies on the lexical boost during syntactic priming (i.e., the observation that syntactic priming is strengthened when words are repeated along with the syntactic structures). This is because our interest is in whether repetition priming can occur when words are repeated across sentences, not in whether priming at the lexical level results in priming at other levels. To this pool, we also added other relevant studies that satisfied the inclusion criteria. As a result, we selected 11 papers tagged as either "lag" or "delay", and 6 tagged as "context". The list of papers described in the review can be found in Table 2.1.

Only two papers (Bartolozzi, Jongman, & Meyer, 2021; Jongman & Meyer, 2017) tagged as "attention" met all the inclusion criteria. Given the scarcity of

List of studies by tag	
Lag/Delay	Context
Cave (1997)	Bassili, Smith, and MacLeod (1989)*
Durso and Johnson (1979)*	Besken and Mulligan (2010)
Francis and Sáenz (2007)	Francis, Camacho, and Lara (2014)
Heath et al. (2012b)*	Gibson and Bahrey (2005)*
Heath et al. (2012a)*	Hamburger and Slowiaczek (1998)
Hernandez and Reyes (2002)	Nicolas and Söderlund (2000)
Mitchell and Brown (1988)*	
Mitchell (1989)*	
Tsuboi, Francis, and Jameson (2021)	
Vitkovitch, Rutter, and Read (2001)	
Wheeldon and Monsell (1992)	

Table 2.1: The table lists the studies tagged as "delay", "lag", or "context" that were discussed in the review. The asterisk (*) indicates that the study was not indexed in the database search but was added from other sources (e.g., reference in other studies).

studies that investigated the effect of divided attention on repetition priming from comprehension to production, a review was not warranted. We therefore only described the two relevant studies in the concluding section and suggested ideas for further research.

2.3 The longevity of priming effects

First, we examined what the effect of lag is on repetition priming. In particular, our aim was to assess whether repetition benefits can still occur when a word is repeated after some time in a conversation and, if so, whether the repetition benefit at delayed repetition is smaller than that at immediate repetition. We discuss the potential role of both intervening items and time delay.

In traditional paradigms where primes and targets are presented in different blocks priming usually occurs, even though prime and target do not immediately follow each other. For instance, in a study by Damian, Dorjee, and Stadthagen-Gonzalez (2011) participants produced nouns in response to definitions in the study phase and, after a short break, they carried out a picture-naming task, in which some of the nouns of the study phase reoccurred as picture names. The robust repetition priming effect indicates that priming can last over a delay (see also Barry et al. (2001)[1] for a similar study employing both reading and picture-naming tasks).

[1] While Barry et al. (2001) and Damian et al. (2011) are mentioned in this section, these studies do not specifically investigate how lag affects repetition priming and are therefore not included in the review.

While it is clear that priming can occur after a delay, a more interesting question concerns the rate of decay of the priming effect. In early priming studies, pictures were repeated at study and test and participants were asked to perform the same task, i.e., picture naming. Mitchell and Brown (1988) tested both repetition priming and recognition memory of pictures that were repeated after one week (Experiment 1), four weeks (Experiment 2), or six weeks (Experiment 3). Repetition priming occurred at all lags, with no differences in the magnitude of the effect.

Mitchell (1989) asked participants to name pictures, some of which were repeated after a lag of 5, 25, or 50 items. The pictures had high or low codability (i.e., name agreement). The data showed that participants named repeated items faster than non-repeated items and that priming effects were stronger for low-codability than high-codability items. As expected, lag modulated latencies for repeated items but not for non-repeated items; in particular, priming decreased at the 50-item lag with respect to the 5 and 25-item lags. In a further analysis on priming scores, the authors confirmed the effects of codability and lag on priming and showed that the decay in priming between 25 and 50-item lags was mainly driven by the low-codability items.

A priming decay across a range of lags similar to the one used by Mitchell (1989) was also observed by Durso and Johnson (1979). In two experiments, they asked participants either to name or categorise a series of words and pictures. The names of the words and pictures could be repeated after a lag (0, 25, or 50 lags in Experiment 1, and 0, 1, 8, or 20 lags in Experiment 2). In the first experiment, priming for named items presented as pictures at both presentations decreased from 200ms at the 0 lag to 165ms at the 25 lag, and to 159ms at the 50 lag. In Experiment 2, facilitation was 109ms at the 0 lag, 126ms at the 1 lag, 131ms at the 8 lag, and 116ms at the 20 lag[2].

In a subsequent study, Cave (1997) had participants perform a picture-naming task in the first session, and a picture-naming task and a recognition task in the second session. The second session took place between 6 and 48 weeks after the first one. Priming was always significant, although the effect decreased across lags. As these studies used picture-naming at both study and test, priming could be due to language-related processes, object identification processes, or both (see Francis et al. (2014)), with the two types of priming effects possibly having different trajectories.

[2]We do not discuss the results of the analysis in detail because these experiments include a task (categorization) and modality of presentation (words presented visually) that do not match the inclusion criteria of this review.

In a classical study by Wheeldon and Monsell (1992) (Experiment 1), participants read words aloud, gave responses to definitions, or named pictures. The pictures were always the targets while the printed words and the words produced in response to the definitions were the primes. The targets could be identical or unrelated to the primes and were separated from the primes by a short lag (2 to 7 intervening trials, corresponding to 10-35s) or a long lag (60 to 120 trials, corresponding to 6-12 minutes). Priming occurred in both conditions but was stronger when the lag was short rather than long. In Experiments 2 and 3, the authors investigated whether picture-naming could be facilitated by the prior production of the prime or a homophone in response to a definition. Primes and targets were separated by a lag between 60 and 120 intervening trials (as in Experiment 1). The authors failed to replicate the long-lag priming effect using homophones at study and test, which led them to conclude that the locus of facilitation was at the interface between semantic and phonological word representations.

Using a similar paradigm to Wheeldon and Monsell (1992), Vitkovitch et al. (2001) had participants respond to definitions (primes) and name pictures (targets). Primes and targets could either be identical, semantically related or unrelated, and were separated by either one lag or three lags. While priming occurred at lag 1, no priming occurred at lag 3 (note that priming effects were significant in the analysis of means but not of medians). The analysis of error rates also showed priming for the lag 1 condition but not for the lag 3 condition. The authors speculated that the absence of priming at the long lag may have been related to the low proportion of identical trials (25%) and/or the higher proportion of semantically related trials (40%).

The effects of lag on repetition priming have also been addressed in two bilingual studies, which are especially useful because they can give us information about the locus of the decay of the priming effect (e.g., articulatory, word retrieval or conceptual level). In Francis and Sáenz (2007), Spanish-English bilinguals either translated words (prime trials) or named pictures (prime and target trials). When the picture name was repeated, participants either named it in the same (non-target language) or different language as when it had first occurred in the translation or picture-naming task. Primes and targets were separated by either a 10-minute or a one-week delay. The analysis of the priming advantages showed main effects of encoding condition and lag (but not of response language), and two two-way interactions: one between encoding condition and response, and one between encoding condition and lag. As for the main effect of encoding condition, priming was stronger when prime and target pictures were named in the same language than in the translation condition; by contrast, the translation condition yielded more priming than picture naming in a different language. The main effect of lag indicated that items in the short lag condition showed more repetition benefits than those in the long lag condition. The interaction between response language and lag showed that priming decayed faster in the non-dominant language. As for the interaction between lag and encoding condition, priming showed a different trajectory across lags according to the task used at encoding. Indeed, picture naming in a different language yielded comparable priming effects at both the short and long lag (response latencies: 1097ms (L1) and 1254ms (L2) vs 1103ms (L1) and 1270ms (L2)). By contrast, priming effects for items in the identical naming condition and in the translation condition were greater when the lag was 10 minutes than one week (response latencies in the identical naming condition: 971ms (L1) and 1010ms (L2) vs 1030ms (L1) and 1117ms (L2); response latencies in the translation condition: 1053ms (L1) and 1118ms (L2) vs 1092ms (L1) and 1212ms (L2)). The attenuation after one week was also replicated in a second experiment where primes and targets were either both translated or named in response to a picture. The fact that lag only affected trials in which prime and target were named in the same language led the authors to conclude that lag affects word retrieval processes.

While Francis and Sáenz (2007) showed that repetition priming for words produced at encoding was evident at both a 10-minute and at one-week delay, a subsequent study from the same lab (Experiment 2; Tsuboi et al. (2021)) showed that word comprehension tasks in which participants read or heard words elicited priming in a picture-naming task at a 10-min delay - as indexed by shorter latencies for repeated versus non repeated words - but not at a one-week delay. Moreover, the priming effect was greater for low-frequency than high-frequency words. The lack of an effect did not depend on whether the encoding task required an overt response (reading or repeating the word aloud) or not, and could therefore not be attributed to differences in phonological and/or articulatory processing between the production tasks in Francis and Sáenz (2007) and in Tsuboi et al. (2021). By contrast, Tsuboi et al. (2021) proposed that conceptual access is not completed in comprehension tasks such as those used in their encoding phase, therefore leading to a faster priming decay.

In two bilingual experiments[3] by Hernandez and Reyes (2002), participants were asked to name pictures, which could be repeated or not. The lag between

[3]Experiment 1 is not reported because it only tested priming at immediate repetition.

prime and target pictures could be 0 or long (36-93 trials). In Experiment 2, the language in which pictures were named (English or Spanish) was consistent within blocks, while in Experiment 3 the language varied within blocks.

When repetition priming was tested within languages, naming latencies decreased after repetition at both lags, more so at lag 0 than at the long lag. By contrast, results from the between-languages experiment showed that repetition benefits did not decrease with lag. The finding that lag affected priming only when the same word - but not its translation - was used at study and test corroborated Francis and Sáenz (2007)'s hypothesis that priming decrements occur at the word retrieval level.

Information about the differences between short and long lag priming also comes two from neuroimaging studies (but see also Van Turennout, Bielamowicz, and Martin (2003); van Turennout, Ellmore, and Martin (2000) for studies using covert picture-naming tasks). In two experiments, Heath and colleagues had participants encode pictures while performing a semantic judgement task on the pictures (Heath et al., 2012b) or while repeating their names, which were presented auditorily together with the pictures (Heath et al., 2012a). Subsequently, participants were asked to name old and new pictures after a lag of a few minutes (about 6-12 lags) or of a few days (with up to two days between the two sessions). Heath et al. (2012b) observed that priming at the shorter lag was mainly associated with increased activation in the left inferior occipital gyrus and in the fusiform gyrus, areas that are usually linked to object identification. By contrast, long-term priming correlated with activation decreases in parts of the middle temporal gyrus that are usually associated with semantic processing and lexical selection. Heath et al. (2012a) found decreased activation in the inferior occipitotemporal areas for items repeated after the short lag, suggesting that repetition priming effects were mainly driven by object recognition and lexical selection processes. By contrast, long-term priming was associated with a decrease in the middle and superior temporal gyri, subserving semantic and phonological processes. Together, the results by Heath et al. (2012b) and Heath et al. (2012a) suggest that, while short-term repetition priming is the consequence of explicit memory contributions and improved object identification processes, long-term priming is compatible with a learning mechanism targeting lexico-semantic processes.

In sum, from the review of the studies above two main findings emerge: first, that repetition priming can occur after a delay (Damian et al., 2011; Francis & Sáenz, 2007; Hernandez & Reyes, 2002; Wheeldon & Monsell, 1992), and second, that priming at longer lags is usually weaker than priming at shorter lags. However, this decrement only occurs when the same word form is used at study and test (Francis & Sáenz, 2007; Vitkovitch et al., 2001), but not when a different word form (e.g., translation of word into another language) is used.

The bilingual studies by Francis and Sáenz (2007) and Hernandez and Reyes (2002) suggest that the loss of facilitation at long lags over short lags is located at the word form level. The neuroimaging studies by Heath et al. (2012b) and Heath et al. (2012a) provide a more complex picture, whereby different circuits subserve short-term and long-term priming. It must be pointed out that these studies are difficult to compare given the differences in the tasks used at encoding (e.g., semantic judgement task, (Heath et al., 2012b), auditory repetition task, (Heath et al., 2012a), picture naming (Francis & Sáenz, 2007; Hernandez & Reyes, 2002); word translation (Francis & Sáenz, 2007)) and in the lags used (0 or between 36-93 trials in Hernandez and Reyes (2002), 10 minutes or one week in Francis and Sáenz (2007), a few minutes or days in Heath et al. (2012b) and Heath et al. (2012a)). While it appears that the locus of priming in language production depends on word retrieval processes, it is not clear yet how these processes are impacted by lag. In the next section, we detail a possible mechanism underlying priming at immediate and delayed repetition.

2.3.1 Explaining lag effects

The main finding emerging from the review of studies on lag is that repetition priming can still occur after a delay but that the amount of facilitation progressively decreases. In this section, we outline possible reasons for this progressive loss of facilitation.

Loss of facilitation can be easily explained if repetition priming is considered as the consequence of a learning-based mechanism (e.g., Howard, Nickels, Coltheart, and Cole-Virtue (2006); Oppenheim, Dell, and Schwartz (2010); but see also Hughes and Schnur (2017)) whereby the connections between concepts and word forms are continuously adjusted to facilitate lexical selection (Oppenheim et al., 2010). According to Oppenheim et al. (2010), processing of a word is associated with the reinforcement of the weights of the semantic-to-lexical connections. This leads to facilitated word retrieval and, therefore, priming, when the prime word is presented a second time. When the connection weights of the target word are strengthened, the association weights of its competitors are weakened. While Howard et al. (2006) do not posit any weakening mechanisms for the competitors, they assume that lexical selection is competitive and that the activation of a strong competitor inhibits that of the target. Both models assume that the reinforcement of the connection weights of a word will facilitate its retrieval at the expenses of its competitors.

We can assume that during a priming session weights of different semantic-to-lexical connections are strengthened at each trial. If the connection weights of a word are not reinforced, the priming effect for that word will progressively fade across intervening trials because the weights of other items are increased in the meantime. This means that, across intervening trials, the initial facilitation will progressively diminish, therefore yielding smaller repetition benefits as the number of intervening trials increases. An explanation of priming decay in terms of weakening of the semantic-to-lexical connections is also in agreement with findings from bilingual studies that priming decays occur at the word retrieval level (Francis & Sáenz, 2007; Hernandez & Reyes, 2002). Given that in the reviewed studies repetition decrements occurred even when no semantically related items were present in the experiment, we must posit that unrelated words can also act as competitors (unlike Oppenheim et al. (2010) and Howard et al. (2006)), either because they are bound as a unique episode together with the target or because they share semantic features with the target.

While we have shown that learning-based models can in principle explain priming decrements, a few questions remain. First of all, one important consequence of these models is that repetition priming decrements should not occur if a word is repeated after some time but without any intervening items. However, both these factors, time and intervening items, can have independent effects on priming, as explained in the introduction to this section. For instance, using a lexical decision task with visually presented words, McKone (1998) showed that both these factors have a detrimental effect on priming. Unfortunately, the above-mentioned studies confounded these two factors, and it is not clear whether time delay and lag contribute differently to the decay of priming effects. We will return to this point in the concluding section.

A second question is whether different mechanisms subserve repetition priming at immediate and delayed repetition. A first possibility is that the priming decrease from immediate repetition to a short delay (within a few minutes) is driven by the aforementioned learning-based mechanism and that the same mechanism underlies priming in both cases. This suggestion would be in line with Francis and Sáenz (2007), who showed that repetition priming due to word production mechanisms - but not word identification mechanisms - is subject to decay. However, in that study the lag varied between ten minutes and one week, while there are no data on how priming processes might differ when repetition is either immediate or occurs after a few minutes. While the hypothesis that priming decay is due to quantitative – rather than qualitative – changes cannot therefore be confirmed, evidence in favour of it comes from an ERP/efMRI study on visual object priming by Henson, Rylands, Ross, Vuilleumeir, and Rugg (2004), where participants were asked to determine, for each presented object, whether they fit into a shoebox or not. Objects were repeated immediately (0 intervening trials, about a 2s delay), at a short lag (1 intervening trials, about 4s delay), at a short delay (0 intervening trials, about 4s delay), and at a long lag (more than 40 intervening items, about a 96s delay). While both the ERP and fMRI data showed that the priming effect decayed at increasing distance between prime and target, there was no evidence that different mechanisms subserved priming at different lags.

A second possibility is that immediate repetition and short-term repetition reflect qualitatively different mechanisms. While priming at short delays might just be a consequence of learning, priming at immediate repetition might reflect additional mechanisms. For instance, it might also reflect episodic contributions. This would parallel explanations given to the lexical boost in the syntactic priming domain, whereby repetition of a word within a sentence increases the likelihood that the same structure will be reused, but only at immediate repetition (e.g., Chang et al., 2006; Hartsuiker et al., 2008). While the reviewed neuroimaging studies suggest that this might be the case, the range of delays used by Heath et al. (2012b) and Heath et al. (2012a) varied between a few minutes and a few days. Unfortunately, there are no available comparisons of the range of delays that are of interest in conversation, i.e., from immediate repetition to a delay of a few minutes.

In sum, the evidence so far suggests that priming can last beyond immediate repetition; however, the longevity of the effect is difficult to determine due to the differences in paradigms across studies. Overall, priming effects are stronger at immediate than delayed repetition. This loss of facilitation can depend on incremental learning processes or episodic memory contributions at immediate repetition, or on the combination of these factors. Further research should first of all establish the differential effect of time and intervening items on the size of repetition priming, by replicating McKone (1998)'s results in other modalities. Secondly, further studies should also determine whether time and/or intervening items impact explicit memory, and determine how such mechanisms can be implemented in current models of priming.

2.4 The role of embedding in priming

Let us now turn to whether repetition priming occurs when words are not presented in isolation but are rather embedded in sentences. In particular, we asked the following questions: if priming occurs for primes embedded in sentences, is the magnitude of the effect different from that obtained with single-word primes? Does the type of embedding affect priming?

In four experiments, Hamburger and Slowiaczek (1998) had participants hear a pair of words that were either unrelated (e.g., *plaque - friend*), shared one or more phonemes (e.g., *flap - friend* or *froth – friend*), or were identical (e.g., *friend – friend*). The first word in each pair constituted the minimal context of the second word, i.e., the prime. After the presentation of the pair, participants were asked to repeat the second word in the pair. Repetition priming from the minimal pairs only occurred in the experiments where the proportion of identical prime-target pairs was high (50%) but not when it was low (25%). A fifth experiment in which prime-target pairs were presented visually showed that the proportion of identical prime-target pairs did not affect repetition priming effects, suggesting that the modulation only occurred when the primes were presented auditorily.

Some evidence that repetition benefits can arise from embedded primes also comes from studies that investigated modality effects in priming. In Bassili et al. (1989), participants read or listened to sentences where the prime word was either presented or inferred. Before each trial, participants were presented with one label: SPECIFIC or GENERAL. The former was used before trials in the inferred condition (e.g., *The boat travelled underwater*), and participants had to think of the exact member of the category (e.g., *submarine*). By contrast, the latter was used before trials in the presented condition (e.g., *The submarine travelled underwater*) and participants were prompted to think of the prime category the word belonged to (e.g., *boat*). In a subsequent stem completion task, priming was significant in all conditions but the effect was stronger when the prime words were presented rather than inferred.

The results obtained by Bassili et al. (1989) were confirmed in a subsequent study by Gibson and Bahrey (2005) (Experiment 2), with the same manipulation of presentation modality (reading vs listening) and prime presentation (presented vs inferred). Unlike Bassili et al. (1989), participants were asked to judge whether the prime word in the sentence referred to a specific or generic concept (e.g., *The food was placed in the refrigerator/appliance*). At the end of the task, they carried out a word-fragment completion task, either in the visual or in the auditory modality, and a cued recall task. Half of the participants carried out the two tasks in the same modality (auditory or visual) while the other half carried them out in different modalities (visual-auditory or auditory-visual). Priming effects - indexed as the proportion of fragments completed with the previously presented primes - were reliable in all conditions and modalities, suggesting that repetition priming can occur when the prime is embedded in a sentence.

Priming effects from embedded primes presented auditorily were also found by Nicolas and Söderlund (2000). Two experiments from that study are relevant: in a first experiment, participants read or listened to short texts and were told that they would later be asked questions about their content. After the study task, participants performed a word-stem completion task and a word recognition task, both presented visually. The results of this experiment showed that priming effects occurred regardless of the study modality; however, they were stronger for visually presented than auditorily presented primes. In Experiment 3, the authors investigated whether modality and target frequency had an interactive effect on repetition priming from embedded words. They showed that priming effects occurred for both high- and low-frequency words in the visual modality; by contrast, primes presented auditorily only yielded repetition benefits if they had low frequency.

While these studies can give us some insights on whether repetition priming occurs for embedded words, they do not reveal how the size of the priming effects compares to the effects of primes presented in isolation. This was addressed by Francis et al. (2014). In their study, participants translated words or sentences in the study phase, and translated words or named pictures in the test phase. In Experiment 1, the to-be-translated words and sentences were presented visually, in Experiment 2 they were presented auditorily. Both experiments showed that priming occurred for words translated both in isolation and in context, although the effect was stronger for words translated in isolation. In Experiment 1, priming occurred both when the task was picture-naming and word translation, but the effect was stronger in the latter condition; by contrast, Experiment 2 showed no differences between encoding tasks.

More evidence about the effects of sentential embeddings on priming comes from Besken and Mulligan (2010). In three experiments, participants heard a series of prime words that were either embedded in sentences (Experiments 1, 2, and 3), recorded in isolation (Experiment 1) or extracted from the sentences containing them (Experiments 1, 2, and 3). In the encoding phase, participants were asked to the rate the clarity of each stimulus. In the test phase, an auditory word-fragment completion task (Experiments 1 and 2) or an auditory word-stem completion task (Experiment 3) was presented to measure priming effects. Experiment 1 showed that priming, indexed as the proportion of fragments completed with the prime words, was significant in the two single-word conditions. By contrast, it was not significant for primes embedded in sentences. The same pattern emerged in Experiments 2 and 3 in which, unlike Experiment 1, the context type was manipulated within-participants, and controlled for the difference in distance between prime and target across conditions. In Experiments 4 and 5, Besken and Mulligan (2010) investigated whether the meaningfulness of the passage impacted on priming effects. Results from the word-fragment competition task showed that embedding primes in a series of unrelated words yielded smaller priming effects than presenting words in isolation; Experiment 5 compared the meaningful and not meaningful embedding conditions and showed that they yielded comparable amounts of priming. In sum, Besken and Mulligan (2010) showed that single words yield more priming than embedded words and that the type of embedding - i.e., meaningful versus not meaningful embedding - does not affect the magnitude of the effect.

In sum, the studies above yielded two main findings. First, priming effects can occur for primes embedded in sentences (Bassili et al., 1989; Francis et al., 2014; Gibson & Bahrey, 2005; Nicolas & Söderlund, 2000; Wheeldon & Monsell, 1992), although they are usually smaller than for primes presented in isolation (Francis et al., 2014), and may even be eliminated (Besken & Mulligan, 2010). Note that the type of embeddings used in the reviewed studies varied not only in terms of meaningfulness but also in length (from one sentence in Bassili et al. (1989) and Gibson and Bahrey (2005) to texts of about 100 words in Nicolas and Söderlund (2000)), and type of prime words used (all studies used nouns apart from Besken and Mulligan (2010), where verbs and adjectives were also employed as primes). While Besken and Mulligan (2010) argued that differences in the meaningfulness of the embeddings do not affect priming decrements, variations in terms of length and type of primes across studies may affect the size of the priming effects. For instance, longer texts contain more words (possibly interfering between each other), and therefore yield smaller priming effects than shorter texts. As for the type of target words, different word classes may have different prominence in a text (e.g., nouns are usually more prominent than adjectives), therefore affecting priming.

The second main finding emerging from the studies above is that the meaningfulness of the embedding does not affect the occurrence or magnitude of priming effects when the target is a single word (a finding that had previously been shown in a masked-priming study by Masson and Macleod (2000)). Below we discuss why priming for single words may be reduced when the words are presented in context at encoding.

2.4.1 Explaining context effects

Let us start from the effect of context on repetition priming for single words. Two theories have been proposed to account for the effect of context on single-word priming: namely, the Transfer Appropriate Processing framework (TAP, Iii, Gallo, and Geraci (2002)), and the Distinctiveness Hypothesis (Masson & Macleod, 2000).

The Transfer Appropriate processing was initially conceived to explain dissociations between tests tapping into implicit and explicit processes. According to TAP, the size of repetition priming depends on the similarity between the cognitive operations engaged at study and at test: the more similar they are, the stronger the priming effect. In particular, it assumes that tests tapping into explicit memory processes are mainly conceptual, while tests tapping into implicit memory processes are perceptual in nature: as a result, performance on these tests depends on the type of processing carried out encoding. However, the TAP assumes that both explicit and implicit memory tests have perceptual and conceptual components. As a result, the use of this approach can also highlight dissociations between implicit tests (e.g., Cabeza (1994)) for differences between the types of processing involved in free association and category association tasks).

Unlike the TAP, the Distinctiveness Hypothesis was developed to explain why scrambled and meaningful tests yielded similar amount of priming in a masked word identification task. According to the distinctiveness hypothesis, words in a sentence are processed in relation to the other items in the sentence, and are therefore not perceived as individual items, leading to priming decrements when the primes are later presented a second time as single words.

The results obtained by Besken and Mulligan (2010) (Experiments 1-3) and Francis et al. (2014) can be easily explained by the TAP account: indeed, the primes were either single words or embedded words, while the targets always required a single word response (e.g., word translation, picture-naming, word fragment completion). This means that the study and test tasks had a greater overlap in the single-word than in the embedded-word condition, leading to stronger priming effects for the former. However, Besken and Mulligan (2010) (Experiment 5) found that context reduced priming for spoken words regardless of whether the primes were embedded in a meaningful context or in a string of words. The authors argued that these results can be better explained by Masson and Macleod (2000)'s distinctiveness hypothesis than the TAP framework. They argued that the TAP framework would predict stronger priming effects for words embedded in strings than for words embedded in meaningful contexts. This is because meaningful contexts require additional conceptual processing: indeed, the meanings of each word in the sentence need to be combined to make sense of that set of words. Instead, such conceptual processing is not necessary for unrelated strings. The distinctiveness hypothesis simply posits that *"To the extent that a word is encoded in a way that renders it distinct, or individuated, from other stimuli, larger amounts of priming will be found"* (Besken & Mulligan, 2010, p. 2026). This view is in agreement with the fact that words that stand out in a text show more priming than words that blend well with the text they are embedded in, as shown for low-frequency rather than high-frequency words or for unpredictable words rather than predictable words (e.g., Nicolas and Söderlund (2000); for studies in different modalities see MacLeod (1989); Rommers and Federmeier (2018a)).

In sum, while the TAP cannot explain some of the context effects for single-word targets, the Distinctiveness Hypothesis can provide a good account of how context impacts priming. What is unclear is whether the priming enhancement for words in lists versus words in sentences, and the priming enhancement for words that stand out in a text versus well-integrated words (e.g., low- versus high-frequency and unpredictable versus predictable words) reflect the same mechanism or whether different processes underlie the effect. For instance, Parmentier (2008) had participants carry out an arrow categorization task. Before the presentation of the arrow, participants heard either a standard sound or a deviant sound. Responses to the categorization task were slower after the presentation of the deviant sound. Furthermore, they were slower after incongruent rather than congruent deviants (e.g., slower responses when the deviant was the word "right" and the arrow pointed left). However, there was no effect of congruency for the standard sounds, which led the author to conclude that auditory distraction was influenced by two distinct effects, a novelty effect and a semantic effect. While Parmentier (2008) focused on the effect of attention on distracting stimuli, a similar dual mechanism might explain the effects of context on priming, whereby a novelty effect would explain the difference between words embedded in lists versus sentences, and a semantic effect underlines differences in priming for embedded words that stand out against the context. We return to this point in the concluding section.

2.5 Summary and conclusions

In this review, we described studies that tested repetition priming of spoken words. Our main aim was to pinpoint the circumstances under which repetition priming can occur, and then evaluate whether repetition priming can work as a supporting mechanism during conversation. We believe that this work is important to be able to evaluate models of conversation that highlighted the role of priming as a supporting mechanism. In particular, we reviewed how both lag and/or time and embedding can have an impact on repetition priming effects.

In the section about lag, we showed that repetition benefits due to word form retrieval decay across time and/or lags and that immediate repetition priming is usually stronger than priming at longer lags. However, the reviewed studies did not enable us to determine whether the difference in priming at immediate and delayed repetition was 1) caused by the number of intervening items or the time delay, and 2) driven by qualitatively similar or different mechanisms. One way to answer these questions would be to first determine how time delay and lag impact priming, and then to investigate the mechanisms underlying any differences. A possibility would be to present participants with a spoken prime, followed by a to-be-named picture of the target at a variable SOA (e.g., SOA=1s vs SOA=5s). At both SOAs, a spoken distractor intervenes between prime and target in half of the trials. After determining whether both time delay and intervening items affect priming, subsequent experiments should investigate the processes modulating priming after intervening items and after a time delay.

In the section about context, we showed that repetition priming of words embedded in sentences is generally reduced with respect to priming of single words, and that the meaningfulness of the context does not impact on the magnitude of priming. This pattern of results can be explained by the Distinctiveness hypothesis, which assumes that repetition priming occurs provided that words can be processed as distinct entities. This account is also able to explain the finding that words that stand out in the sentence (e.g., low- versus high-frequency words or congruent versus incongruent words) elicit more priming than words that fit well with the preceding context.

As pointed out in the review, the priming difference obtained for embedded versus single words and words that do not stand out or stand out against the context might be driven by different mechanisms. In particular, we suggested the possibility that the former is a consequence of a novelty effect, whereby single words capture more attention than embedded words; by contrast, the latter is a consequence of a semantic effect, whereby the relationship between the words and the preceding context influences the size of priming. One way to test this hypothesis would be to present participants with identical and unrelated primes preceded by another word, which would therefore constitute a minimal context. In order to determine whether in previous experiments the difference in priming between single and embedded words was driven by novelty, word pairs should either have a constant SOA in a high proportion of trials, and a deviant SOA in a small proportion of the trials. If novelty drives differences in priming, repetition priming should be greater for the words in the subset with a deviant SOA than for words with a standard SOA. In addition to this novelty effect, a follow-up study could also investigate any additional effects of semantic congruency. For instance, the target could be preceding by a congruent or incongruent adjective (e.g., *wooden chair* versus *blond chair*). If congruency drives the effect, priming should be greater for the trials where the target does not fit with the preceding context, and the effect should be stronger for trials including a deviant SOA.

So far, we outlined possible mechanisms underlying the influence of lag and context on repetition priming. Investigating these mechanisms is important not only for theoretical purposes, i.e., models for repetition priming, but also for practical purposes, that is to say to determine whether and how repetition priming can support conversation and how repetition priming interacts with other linguistic phenomena.

For instance, it has been widely shown that repeated words enhance syntactic priming, a phenomenon known as lexical boost. Unlike syntactic priming, the lexical boost dissipates after a few intervening trials (Hartsuiker et al., 2008; Mahowald, James, Futrell, & Gibson, 2016). Investigating the mechanisms responsible for immediate versus short-term priming on the one hand, and priming for single versus embedded words on the other hand could therefore help explain how repetition priming facilitates priming at other levels (e.g., syntactic) and why any decays occur.

In addition to further exploring the relationship between repetition priming and syntactic priming, further studies should investigate whether different types of embeddings affect priming and whether the type of embedding affects priming decays differently across lags and time delay. For instance, it would be interesting to consider the relationship between repetition priming and predictability:

while unpredictable words are usually associated with a processing cost (e.g., DeLong, Urbach, & Kutas, 2005), there is some evidence that they yield more priming than predictable words (e.g., Rommers & Federmeier, 2018a). Similarly, future studies could further investigate whether repetition priming is affected by the position of the prime word in the sentence, (e.g., in focus or out of focus; Camblin, Ledoux, Boudewyn, Gordon, & Swaab, 2007).

One final issue that we did not discuss in the review is the effect of attention on repetition priming. Exploring this issue is important because research on turn-taking suggests that people might sometimes dual task between production and comprehension (Bögels et al., 2015; Corps et al., 2018; Fargier & Laganaro, 2016). However, linguistic dual-tasking is often associated with performance decrements (Bögels et al., 2018; Fargier & Laganaro, 2019). It is unclear whether repetition priming can support conversation even when participants dual task between production and comprehension.

Two studies (Bartolozzi et al., 2021; Jongman & Meyer, 2017) have shown that priming for spoken words encoded during speech planning does not differ from repetition priming for words encoded in silence. While these studies suggest that priming is resilient to divisions of attention, comprehension studies about priming suggest that repetition priming still requires some degree of attention (e.g., Wood, Stadler, & Cowan, 1997). It might therefore be the case that the paradigms used by Jongman and Meyer (2017) and Bartolozzi et al. (2021) were not sensitive enough to show any effects of divided attention on priming. For instance, in both Jongman and Meyer (2017) and Bartolozzi et al. (2021) prime and target were separated by a relatively long interval, at least 2s, which means that they might have shifted their attention from the distractor task to the main task. By contrast, effects of attention on priming might be evident by, e.g., increasing the difficulty of the distractor task or increasing the frequency of the response to the distractor task, as shown in comprehension studies (Wood et al., 1997, e.g.,) and studies using non-linguistic distractor tasks (e.g. Mulligan, Duke, and Cooper (2007)).

2.6 Conclusions

In this review, we described previous experimental studies that tested repetition priming from either comprehension or production to production, with the aim of describing the situations in which priming could aid conversation. We described how lag, context, and, to a more limited extent, attention influence priming and outlined suggestions about how future studies could explain the effect of these factors on repetition benefits and how the role of repetition priming in conversation could be further explored.

3 | Concurrent speech planning does not eliminate repetition priming from spoken words: Evidence from linguistic dual-tasking

Abstract

In conversation, production and comprehension processes may overlap, causing interference. In three experiments, we investigated whether repetition priming can work as a supporting device, reducing costs associated with linguistic dual-tasking. Experiment 1 established the rate of decay of repetition priming from spoken words to picture naming for primes embedded in sentences. Experiments 2 and 3 investigated whether the rate of decay was faster when participants comprehended the prime while planning to name unrelated pictures. In all experiments, the primed picture followed the sentence featuring the prime on the same trial, or ten or fifty trials later. The results of the three experiments were strikingly similar: robust repetition priming was observed when the primed picture followed the prime sentence. Thus, repetition priming was observed even when the primes were processed while the participants prepared an unrelated spoken utterance. Priming might therefore support utterance planning in conversation, where speakers routinely listen while planning their utterances.

Keywords: repetition priming, speech comprehension, speech planning, divided attention

3.1 Introduction

Holding a conversation seems an effortless task, yet it requires tight coordination between the speakers. Indeed, analyses of corpora of conversational speech suggest that gaps between turns are often only around 200 to 300ms in duration (Levinson & Torreira, 2015). By contrast, laboratory studies of word production, usually for picture naming, report latencies of at least 600ms (Indefrey & Levelt, 2004), and sentence production latencies often exceed a second (Allum & Wheeldon, 2007; M. Smith & Wheeldon, 1999). Some studies have suggested that short gaps in conversation arise because next-turn speakers start planning a response before the end of their interlocutor's turn (Levinson & Torreira, 2015). This means that speech planning and comprehension might overlap in time, a process that we will refer to as linguistic dual-tasking. Early planning may support fast turn-taking, but it should impose a substantial cognitive load because comprehension and speech planning processes must be performed simultaneously. It is unclear how interlocutors deal with these processing costs. In this study we explore whether the burden of linguistic dual-tasking can be reduced by repetition priming, which has already been identified as a pivotal mechanism in models of conversation (Pickering & Garrod, 2004, 2013).

3.1.1 Linguistic dual-tasking in conversation

Previous experimental work has shown that speakers often start planning responses, while still listening to their interlocutors (Barthel, Meyer, & Levinson, 2017; Barthel, Sauppe, Levinson, & Meyer, 2016; Bögels et al., 2018, 2015; Corps et al., 2018; Lindsay, Gambi, & Rabagliati, 2019; Magyari, De Ruiter, & Levinson, 2017). These studies have also revealed that such linguistic dual-tasking reduces turn gaps, but that speech planning is less efficient than when it occurs in silence (Fairs, Bögels, & Meyer, 2018; Fargier & Laganaro, 2016, 2019). For instance, in a study by Bögels et al. (2015) participants answered quiz-style questions, such as *"Which character, also called **007**, appears in the famous movies?"*. There were two experimental conditions. In the early cue condition, the cue to the answer (**007** in the example) appeared in the middle of the question, whereas in the late cue condition, it occurred at the very end, as in *"Which character from the famous movies, is also called **007**?"*. The participants were asked to answer as fast as possible. Speech onset latencies measured from the end of the question were shorter in the early cue (640ms) than in the late cue condition (950ms), indicating that in the early cue condition participants started planning their response during the question. Yet, the response time advantage (310ms) was much less than the time between the cues in the two conditions (1707ms). This means that the participants' response planning was less efficient in the early cue than in the late cue condition. If planning had been equally efficient in both conditions, responses in the early cue condition should have started well before, rather than after the end of the question.

The results of the study by Bögels et al. (2015) are consistent with numerous word production studies that have shown, first, that speech planning requires attention and therefore suffers when carried out concurrently with another task that also requires attention (e.g., Almor, 2008; Boiteau et al., 2014; Sjerps & Meyer, 2015) and, second, that interference arises when linguistic representations are activated simultaneously for speech comprehension and production. For instance, in dual-task experiments, there is more mutual interference when picture naming is combined with syllable- than with tone-monitoring (Fairs et al., 2018; Fargier & Laganaro, 2016). This indicates that the spoken distractors are processed and affect speech planning.

The experimental work on linguistic dual-tasking suggests that speaking in conversation should be rather effortful. However, these experiments generally employed unrelated production and comprehension stimuli, whereas turns in conversation often have some degree of coherence and refer to the same topic. This means that interlocutors can refer to concepts introduced earlier and use words that have occurred before to refer to relevant entities or events. Thus, they might reduce the linguistic dual-tasking costs by priming each other. The question addressed in the present paper was whether such priming could occur when participants were engaged in linguistic dual-tasking, as they often are in natural conversation.

3.1.2 Two roles of priming in conversation

Priming might aid conversation in two related ways, by affecting the choice of words and thereby supporting the creation of common ground between the speakers, and by increasing the speed of comprehension and production processes. Common ground is background knowledge that is shared by the interlocutors and is used to shape and guide conversations (e.g., Clark & Marshall, 1981). Establishing common ground entails, amongst other things, that speakers agree on names for referents in common ground (e.g., consistently talking about "the shoe" or "the trainer"). Such agreement renders utterances progressively less ambiguous by establishing single referents for words that could refer to multiple referents, therefore making them easier to understand. It may also facilitate speech planning by supporting the appropriate choice among alternative ways of referring to referents.

There is a large body of work concerning the processes underlying the establishment of common ground (Arnold, 2016; Brown-Schmidt, 2012; Brown-Schmidt & Duff, 2016; Clark & Marshall, 1978; Duff, Hengst, Tranel, & Cohen, 2006; Horton, 2005, 2007; Horton & Gerrig, 2016). The influential model of dialogue proposed by Pickering and Garrod (Garrod & Pickering, 2009; Pickering & Garrod, 2004) highlights the role of priming for the alignment of the interlocutors' situation models; i.e., their representations of the situation under consideration. The proposal is that the establishment of shared situation models (called "implicit common ground" by Garrod and Pickering) is strongly driven by "a priming mechanism, whereby encountering an utterance that activates a particular representation makes it more likely that the person will subsequently produce an utterance that uses that representation" (Pickering and Garrod (2004), p. 173). This model posits that priming is an automatic and unconscious process, which cascades through the levels of the language system. In other words, priming can arise at one level (e.g., the lexical level) and spread to other levels (e.g., the syntactic and phonological level), so that interlocutors become fully aligned. Other authors have argued that automatic priming does not suffice to explain the establishment of common ground (e.g., Brown-Schmidt & Duff, 2016), but that other, more deliberate processes are also involved. Nonetheless automatic repetition priming is generally seen as one of the mechanisms supporting the establishment of common ground by increasing the likelihood that interlocutors converge on a joint vocabulary to refer to the concepts under consideration.

Priming can affect both the lexical choice and the speed of speech planning, therefore contributing to swift turn-taking. It appears that this potential function of priming in conversation has not been discussed much in the literature, but it is central to the current research. In the following section we review studies of repetition priming, which do not directly concern conversational turn-taking but provide information about the conditions under which priming may be expected to occur.

3.1.3 Repetition priming of word production

There is a large literature showing that recent exposure to a word can facilitate later comprehension or production of the same word (Francis et al., 2014; McKone, 1995; McKone & Dennis, 2000; Monsell et al., 1992; Wheeldon & Monsell, 1992). Many studies of repetition priming have assessed the effects of written or spoken word primes on the subsequent processing of other written or spoken words (e.g., Forster & Davis, 1984; Versace & Nevers, 2003), but in the present context studies of primed word production are more relevant. In these studies, participants typically name target pictures. A variety of priming tasks have been used including picture naming, definition naming, and translation (e.g., Barry et al., 2001; Brown, Neblett, Jones, & Mitchell, 1991; Durso & Johnson, 1979; Francis et al., 2014; Heath et al., 2012b; Johnson, Paivio, & Clark, 1996; Mitchell & Brown, 1988). These studies have shown that repetition can facilitate both object recognition and the retrieval of object names from the mental lexicon. Different proposals have been made about the way repetition might facilitate lexical access; most commonly it is assumed that the repetition benefit arises at the semantic-phonological interface (Monsell et al., 1992; Wheeldon & Monsell, 1992). As repetition priming effects on picture naming can be observed across long lags between primes and targets, it is often seen as a form of implicit learning, involving long-lasting changes of the activation levels of processing units or their connections (Hughes & Schnur, 2017; Monsell et al., 1992; Oppenheim et al., 2010).

Most studies of repetition priming in picture naming used tasks where the participants produced the prime words as well as the targets. However, since the linguistic representations involved in speaking and listening are largely shared or tightly linked (e.g., McQueen and Meyer (2019) for discussion), repetition should also be observed when participants hear prime words and produce picture names. Research from different lines of research support this prediction.

First, picture-word interference studies, where participants name pictures while listening to distractor words, have sometimes included an identity condition, where the distractor corresponded to the picture name (e.g., Glaser and Düngelhoff (1984); Schriefers, Meyer, and Levelt (1990); see Costa, Miozzo, and Caramazza (1999) for a bilingual study). Compared to unrelated and neutral conditions (featuring non-words or noise as distractors) identity primes yield facilitation, pointing to a repetition benefit.

A second relevant line of work is research on the lexical boost in structural priming. Structural priming is the observation that speakers become more likely and sometimes faster to use certain grammatical structures (e.g., passives) after exposure to these structures (e.g., J. K. Bock, 1986; K. Bock, Dell, Chang, & Onishi, 2007; Dell & Ferreira, 2016). Structural priming effects can be boosted when prime and target sentences share content words (e.g., Cleland & Pickering, 2003; Segaert, Kempen, Petersson, & Hagoort, 2013; Segaert, Wheeldon, & Hagoort, 2016). This indicates repetition priming from words embedded in the priming sentences. The lexical boost effect has been found to be more short-lived than the structural priming effect and has therefore been proposed to be an episodic memory effect (Hartsuiker et al., 2008).

Third, there is a substantial body of work that is related to the work on common ground mentioned earlier and concerns the way utterances are produced when they refer to novel concepts versus concepts introduced earlier. The key observation is that utterances that refer to concepts mentioned earlier are reduced in duration, specifically those words that have occurred before (e.g., Bard, Aylett, Trueswell, & Tanenhaus, 2004; Jacobs, Yiu, Watson, & Dell, 2015; Kahn & Arnold, 2015). A widely discussed issue is whether reductions serve audience design, i.e., occur for the benefit of the listener, or result from speaker-internal processes. Most important for the present discussion is the observation that repeated words are sometimes not only reduced in duration but also initiated earlier. For instance, Kahn and Arnold (2015; Experiment 2B) had two participants (a speaker and a listener) perform a joint task. One of them saw an event on their screen and had to instruct the other participant to create the same event on their screen using utterances such as *Make the [object] flash*. The speaker would then instruct the listener to move another object (identical or different than the previous one). The speaker initiated this utterance faster and reduced their duration when the object name was repeated rather than novel, regardless of whether they or their partner had just produced the object name. These findings indicate repetition priming, i.e., facilitation of speaker-internal processes from heard words onto speech production.

In sum, there is strong evidence that word production may be speeded by repetition priming from words heard earlier. In the following sections we consider three factors that may limit the strength of such repetition priming effects in conversation: the lag between primes and targets, the embedding of primes in sentence contexts, and linguistic dual-tasking. These factors were assessed in an earlier study in our lab (Jongman & Meyer, 2017), which will be described below, as well as in the present study.

3.1.4 Limiting conditions for repetition priming

Turning first to the lag between primes and targets, many studies have highlighted the longevity of repetition priming effects. For instance, repetition benefits for picture naming have been reported after several days or even years (Cave, 1997; Mitchell, 2006; Van Turennout et al., 2003). Other studies have varied the lag between prime and target within experimental sessions (e.g., Durso & Johnson, 1979; Wheeldon & Monsell, 1992). Repetition priming effects tend to decay over time and/or with the amount of intervening materials, but are nonetheless still measurable after considerable delays. Durso and Johnson (1979) reported repetition priming effect for lags up to 50 intervening items. Wheeldon and Monsell (1992) reported a repetition priming effect for lags ranging between 2 and 7 items (10-35 seconds, short lag condition) and a weaker effect for lags ranging between 60-120 items (6-12 minutes, long lag condition). However, in all of the picture naming studies, participants produced the primes, either in response to definitions or to pictures. Thus, there appears to be no evidence about the longevity of repetition priming from heard words to picture names. Given that repetition priming effects are generally weaker from heard than produced primes, it is not obvious that priming effects from heard primes will be maintained over longer lags. Consistent with this suggestion, the lexical boost effect described above has been characterized as short-lived.

Next we consider the effect of embedding primes in sentence contexts. Sentential embedding of primes has often been studied in paradigms involving target comprehension (Coane & Balota, 2010; Levy & Kirsner, 1989; MacLeod, 1989; Masson & Macleod, 2000; Oliphant, 1983; Speelman, Simpson, & Kirsner, 2002). A robust finding is that sentential embedding reduces and sometimes eliminates repetition priming effects. A word production study using embedded primes was conducted by Francis et al. (2014). In the priming phase of this study, bilingual participants translated words presented either in sentence context or in isolation. In the test phase, they were asked to translate words or to name pictures. As in the comprehension studies, the priming effect was reduced when words were initially translated in context rather than in isolation. An account of the effects of contextual embedding on repetition priming is that embedding prime words in sentences or, in fact, lists of words, affects the way these items are encoded (Masson & Macleod, 2000). According to this view, the distinctiveness account, priming effects are strongest when the primes are *"distinctively encoded and individuated against the background of other items that are presented"* (p. 1096).

None of the studies on the effects of presenting primes in context has involved the task combination at issue here, namely listening to primes and producing target words. However, the distinctiveness account, if valid, should apply here as well. Consequently, one would expect weaker repetition priming from contextually embedded spoken prime words onto word production than from isolated prime words. Recall though that priming effects from heard primes embedded in sentences were found in the research on the syntactic boost effect and on audience design described above.

Finally, we consider whether repetition priming can occur when speakers are planning utterances while hearing the primes. This is important because, as indicated above, speakers often engage in planning while listening to their interlocutor, dividing attention between comprehension and production processes. Many studies of verbal memory have assessed the effects of divided compared to full attention to word processing on later memory. As one might expect, the effects depend on many variables, including the type and difficulty of the secondary task and the way memory is assessed, in particular through explicit memory (recall or recognition of the words) or implicit memory tasks (most commonly word fragment completion, word association, category-exemplar generation; e.g., Gabrieli et al. (1999); Mulligan and Stone (1999); Wolters and Prinsen (1997); for a review see Spataro, Cestari, and Rossi-Arnaud (2011)). The emerging pattern is that divided attention can reduce priming effects when the secondary task is difficult (Mulligan, 1997, 1998). Furthermore, the results suggest that at least some attention to the prime is needed for priming effects to occur (e.g., Keane, Cruz, & Verfaellie, 2015; MacDonald & MacLeod, 1998; Mulligan, 1998).

Repetition during linguistic dual-tasking has so far only been investigated by Jongman and Meyer (2017). On each trial of this study participants heard a prime word and simultaneously saw a distractor picture followed by a target picture. The prime was identical to the name of the target, associatively related to it, or unrelated. In the no-plan condition, participants were asked to name only the target picture; in the plan condition, they had to name both the distractor and the target picture. Condition was manipulated between participants in Experiment 1 and within participants in Experiment 2. While associative priming was absent in the plan condition of Experiment 2, robust repetition priming was found in the plan and no-plan conditions of both experiments. Thus, there was repetition priming from hearing prime words to producing them as picture names, even during dual-tasking.

In sum, the existing literature shows that repetition priming can arise when prime and target are separated by intervening items, when the primes appear in utterance contexts, and when participants engage in other tasks, including speech planning, while processing the primes. However, no study has investigated whether repetition priming is seen when these conditions are met simultaneously, as will often be the case in conversation. That is, it is not known whether repetition priming from comprehending a prime word to producing that word occurs when the prime is embedded in an utterance context, when there is a lag between prime and target, and when the participant is preparing an utterance while hearing the prime. The goal of the present study was to answer this question. The results should contribute to a better understanding of the cognitive processes underlying repetition priming. More importantly, they should provide new evidence about the potential importance of repetition priming for alignment and swift turn-taking in conversation.

3.2 The current study

The paradigm used in the present study was similar to the paradigm used by Jongman and Meyer (2017). The participants heard prime words and produced target words as picture names. However, in contrast to the earlier study, the primes were embedded in sentence contexts. On the trials of Experiment 1, participants listened to a sentence containing a prime word (e.g., *Hij heeft helaas de radio kapotgemaakt*, Unfortunately he has broken the radio) and saw two pictures one after the other, Picture 1 and Picture 2, which they had to name (see Figure 3.1). The prime word could be identical or unrelated to Picture 2 (the target), and was always unrelated to Picture 1. Picture 1 appeared after the offset of the sentence (i.e., in silence), 2s after sentence onset. The prime sentence and Picture 2 appeared on the same trial (no lag condition), or were separated by 10 or 50 intervening trials (short lag and long lag condition, respectively). Picture 2 in the short lag condition occurred one minute after the prime, Picture 2 in the long lag condition occurred five minutes after the prime.

This experiment studied whether repetition priming from word comprehension to picture naming occurs when the primes are embedded in a sentence context, and explored how any priming effects would change when prime and target follow each other immediately or are separated in time and by intervening materials. As reviewed above, the relevant literature does not reveal how the effects of contextual embedding and lag might combine for the types of prime and target stimuli at issue here. Our prediction was that the priming effect should be strong (possibly equivalent to the effect obtained by Jongman and Meyer (2017) for word primes) in the immediate condition, but should decrease across the lags.

Experiment 2 used the same paradigm in order to determine whether priming was affected by division of attention between comprehension and concurrent speech-planning. Unlike Experiment 1, Picture 1 now appeared at the onset of the prime word in the sentence. Therefore, participants initiated speech planning while still comprehending the sentence. In the study by Jongman and Meyer (2017), the repetition priming effect was unaffected by division of attention. Here we studied whether this was also the case when the primes were embedded in sentence contexts and when prime and target were separated in time. If the amount of attentional resources allocated to the comprehension task in Experiment 2 allows enough processing of the stimulus, the magnitude and the rate of decay of the priming effect should be similar in Experiments 1 and 2. By contrast, if splitting attention between the comprehension and production taskS does not enable thorough processing of the auditory stimulus, the rate of decay of the priming effect should be greater in Experiment 2 than in Experiment 1.

Any priming effects obtained in Experiment 2 might be influenced by task demands. Given that the relationship between prime and Picture 2 in the no lag condition was quite obvious, participants might strategically pay attention to the auditory input because it may support subsequent picture naming. For this reason, we also carried out a modified version of Experiment 2 in which the no lag condition was removed (Experiment 3).

In the study by Jongman and Meyer, prime words and target pictures were repeated twelve times during the experiment (six times as prime and six times as target). Some studies have shown that prime and/or target repetition may increase the size of priming effects (Jacoby & Dallas, 1981). In the present study, primes and targets were only presented once, thereby allowing us to assess whether repetition priming would be obtained when the experimental items were not repeated.

3.3 Experiment 1

In the first experiment, participants heard sentences containing a prime word and were subsequently asked to name first Picture 1 and then Picture 2. The name of Picture 2 was either identical or unrelated to the prime word. In the no lag condition Pictures 2 appeared on the same trial as the prime word, in the short lag condition they appeared 10 trials (1 minute) after the prime word, and in the long lag condition, they appeared 50 trials (5 minutes) after the prime word. We refer to the conditions as no, short or long lag conditions, but it is important to point out that time and number of intervening items may differentially contribute to the rate of decay of the priming effect (McKone, 1998). Picture 1 was presented after the prime sentence had ended, ensuring no overlap between the presentation of the prime and Picture 1. The main aim of Experiment 1 was to replicate previous studies that found decay of the priming effect as a function of time/number of intervening trials, with the novel addition of embedding the prime in a sentence context. This configuration was intended to mimic situations occurring in conversation where a word in one speaker's turn might prime the next speaker's use of the same word.

3.3.1 Method

Power analysis

In order to determine whether we would be able to detect an effect of Condition on Picture 2 latencies, we carried out a power analysis. Condition was modelled using Helmert contrasts, as in the main experiments. The first contrast compared the unrelated condition against all the other conditions, the second contrast compared the long lag condition against the short lag and no lag conditions, and the third contrast compared the short lag condition against the no lag condition. In the simulation (n=1000), the significance of each contrast was tested by comparing the model that included all contrasts with the model that did not include the contrast of interest. The mean reaction times for each condition for the power analysis were taken from Experiment 1 in Jongman and Meyer (2017), which offers the best approximation to our conditions and design. In their experiment, the mean naming latency in the identical priming condition was 597ms and mean naming latency in the unrelated condition was 679ms. The estimates for the short and long lag condition were chosen on the basis of the results obtained by Wheeldon and Monsell (1992), who tested word-to-picture priming at longer delays. In their experiment, the mean naming latency in the long lag condition (prime and target separated by 60-120 lags, 6-12 minutes) increased by 4.73% with respect to the short lag condition (prime and target separated by 2-7 lags, 10-35ms). The estimates in Wheeldon and Monsell (1992) for the short lag and unrelated condition, 592ms and 688ms respectively, were similar to those obtained for the immediate and unrelated condition in Jongman and Meyer (2017). Therefore, we first calculated the percentage of mean latency increase from the short lag to the long lag condition in Wheeldon and Monsell

(1992). We then applied this percentage increase to Jongman and Meyer's identical condition, so as to build a mean naming latency for our short and long lag conditions. As a result, in our power analysis we used a mean naming latency of 625ms for both the short and long lag conditions.

Assuming 160 items and 40 participants, the simulated data yielded a power of 0.91 at α =.05. The power analysis and the raw data of each of the experiments presented in this paper are available on the MPI for Psycholinguistics Archive (https://archive.mpi.nl/mpi/).

Participants

Forty participants (8 male, mean age: 22.65 years, range: 18-27 years) took part in Experiment 1. They were recruited from the Max Planck Participant Database, were native speakers of Dutch, did not report any speech problems, and had normal or corrected-to-normal vision. Participants received € 8 as compensation for taking part in the study. Ethical approval for all experiments reported here was granted by the Social Sciences Faculty of Radboud University.

Design

The experiment included 290 trials. During each trial, participants listened to a sentence containing a prime word (e.g., *Hij heeft helaas de **radio** kapotgemaakt*, Unfortunately he has broken the radio) and then saw two pictures, Picture 1 and Picture 2, which they had to name (see Figure 3.1). In Experiment 1, Picture 1 was presented at the end of the sentence, allowing participants to process the prime in silence. Picture 2 was presented 2s after the onset of Picture 1.

130 of the 290 trials were filler trials. Fifty of the filler trials were presented at the beginning of the experiment to allow for an even spread of long lag trials across the experiment, rather than presenting them all towards the end. These filler items did not include prime words. In the remaining 80 filler trials, the sentence contained a prime word that corresponded to a Picture 2 name on a different trial (i.e., the prime sentence in the lag conditions, see below). Responses to pictures on filler trials were not analyzed.

160 trials were experimental trials. Picture 1 was always unrelated to the prime word in the trial. Picture 2 was always a target picture and, as described above, appeared in one of four different conditions: the no lag condition, short lag condition, long lag condition, and unrelated condition. In the no lag condition, Picture 2 was immediately preceded by a sentence containing its name.

For example, the picture of a radio appeared on the same trial as the sentence *Hij heeft helaas de* ***radio*** *kapotgemaakt*. In the short lag condition, Picture 2 appeared ten trials (corresponding to one minute) after the sentence containing the prime; in the long lag condition, Picture 2 appeared fifty trials (corresponding to 5 minutes) after the sentence containing its prime; in the unrelated condition, the label of Picture 2 did not occur anywhere else in the experiment.

We created four item lists. Each list included all experimental and filler items. Condition was counterbalanced across items. That is, in each list, 40 different Picture 2's appeared in each condition. Each list was presented to ten participants. Participants were randomly assigned to lists.

Materials

Sentences. A female speaker of Dutch produced the sentences with neutral intonation. Recordings were made in a soundproof booth using a Sennheiser ME 64 microphone. Sentences were recorded using the software Audacity (Audacity-Team, 2012). All sentences had a length below 2000ms (average length: 1879ms, range: 1493ms – 2000ms). Two sentences which were initially longer than 2000ms (2054ms and 2060ms) were compressed using Praat (Version 5.1, Boersma and Weenink (2009)). The auditory stimuli were normalised in order to ensure that they had similar intensity. Each sentence included one concrete noun, the prime. The prime was preceded by an adverb and followed by a past participle (e.g., *Hij heeft helaas de radio kapotgemaakt*, "He has unfortunately the radio broken").

Pictures. 307 pictures were taken from the MultiPic database (Duñabeitia et al., 2018). 145 pictures were chosen as Picture 1 and 160 pictures were chosen as Picture 2. Two pictures appeared on practice trials. Repetition of some pictures was necessary because the database did not include enough suitable pictures. Items chosen as Picture 1 had a mean name agreement of 87.65% (range: 38.60% - 100%) and mean frequency (fpm, frequency per million in the SUBTLEX database, Keuleers, Brysbaert, and New (2010)) of 112.87 (range: 0.02 – 4412.02). Some of the items chosen as Picture 1 were repeated across the experiment as Picture 1 in other experimental trials (15 items, repeated twice) or as Picture 1 or Picture 2 in filler trials. These items were repeated for a maximum of three times throughout the experiment. Items chosen as Picture 2 had a mean name agreement of 92.82% (range: 54.39% - 100%) and a mean frequency of 8.83 (range: 0.02 – 29.29). All Picture 2 names were low to medium in frequency (fpm < 30; words with fpm < 5 are traditionally classified as low-frequency words, words with fpm > 100 are traditionally classified as high-frequency words; Brysbaert, Mandera, and Keuleers (2018)). Low- to medium-frequency items were preferred over high-frequency ones because priming effects are usually larger for low- than for high-frequency targets (Wheeldon & Monsell, 1992). Picture 2 items were never repeated.

Procedure

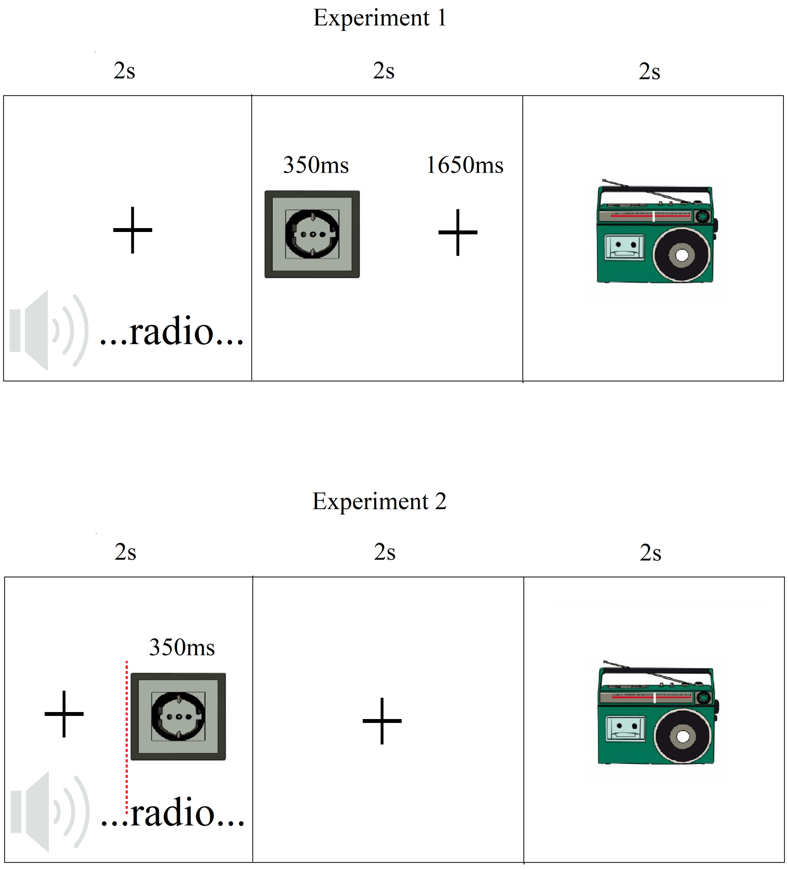

Figure 3.1: Trial structure in Experiments 1 and 2. This example shows a trial in the no lag condition. In each trial, participants hear a prime word (e.g., *radio*), and see two pictures, Picture 1 (e.g., *socket*) and Picture 2 (e.g., *radio*).

All trials had the same structure (see Figure 3.1). At trial onset participants heard a sentence containing a prime word while looking at a fixation cross. Sentence duration was maximally 2s. Two seconds after the onset of the sentence, the fixation cross was substituted by Picture 1, which was shown for 350ms and followed again by a fixation cross. Two seconds after the onset of Picture 1, the fixation cross was replaced by Picture 2. Picture 2 remained on screen for 2s, the second response window. The SOA between the onset of the sentence and the onset of Picture 2 was 4s. At the offset of Picture 2, a new trial began. Each trial lasted 6s in total. Participants were instructed to listen to the sentences and name each picture as soon as it appeared on the screen. The experiment lasted about 40 minutes.

Before the beginning of the experiment, participants carried out a practice session to familiarize them with the task. The practice session included four trials, which were structured as the trials in the actual experiment. We presented two trials in the no lag condition, and two trials in the unrelated condition.

Apparatus

The experiment was controlled using the software Presentation (version 20.0, www.neurobs.com). Sentences were played using headphones (Sennheiser HD 437) and responses were recorded using a Sennheiser ME 64 microphone. Stimuli were presented on a 17” monitor (Iiyama LM 704U7).

Scoring and analysis

Responses to the pictures were coded as correct if participants used the dominant name in the picture database, which, for Picture 2 items, corresponded to the primes. A trial was coded as correct when both Picture 1 and Picture 2 were named correctly, and as incorrect if one of the two pictures or both were named incorrectly. Only correct trials were included in the analyses of response latencies.

Three items, used as Picture 1 in experimental trials or as Picture 1 and 2 in filler trials, were mistakenly presented four times (instead of three times) in some of the lists. Therefore we removed the experimental trials where these items occurred for the fourth time. One experimental item was removed from all lists because it was erroneously repeated. Response latencies were measured manually by the first author using the software Praat (version 5.1, Boersma and Weenink (2009)).

Data from experimental trials were analysed using mixed effects models (package lme4, version 1.1.14, Bates, Mächler, Bolker, and Walker (2014)) in R (R Team, 2017, version 3.4.1). We modelled the rates of accurate responses to Picture 1 and Picture 2 and the naming latencies for both pictures on correct trials (i.e., trials where both pictures were named correctly).

The most important dependent variable was the naming latency for Picture 2. Since the naming latencies were right-skewed, they were log-transformed and then trimmed by participant and condition using a cut-off value of 2.5SD beyond the mean (values outside the cut-off were excluded). Picture 1 latencies were also analysed to make sure that any differences in Picture 2 latencies (and accuracy) were not related to differences in the difficulty of processing Picture 1. All latencies were analysed using linear mixed effects models.

Given that in Jongman and Meyer (2017) the rates of correct responses were high and not affected by the primes, we did not predict any priming effects for the rates of correct responses in the present paper. However, the primes could, of course, also affect accuracy, and to explore this possibility, accuracy rates for Picture 2 were analysed as well. Picture 1 accuracy rates were also analysed. As in the case of Picture 1 latencies, Picture 1 accuracy rates should not depend on the experimental conditions. In both Picture 1 and Picture 2 analyses, Accuracy was a categorical variable (1 = correct trial, 0 = incorrect trial) and was analysed using generalised mixed effects models.

All the models were run using the optimizer BOBYQA (Powell, 2009). Condition was always the independent variable and modelled using Helmert contrasts. The first contrast compared the unrelated condition to the other conditions (long lag, short lag, no lag), the second contrast compared the long lag condition to the short and the no lag condition, and the third contrast compared the short condition to the no lag condition. While Helmert contrasts can give us information on how priming decays across lags, they do not enable us to directly compare whether each of the priming conditions yields repetition benefits with respect to the unrelated condition. In the model, the intercept represents the overall mean.

All models were initially built using a maximal random-effects structure. Accuracy and latency models for Picture 2 initially included by-participant and by-Picture 2 intercepts and slopes for Condition. Accuracy and latency models for Picture 1 initially included by-participant and by-Picture 1 intercepts and slopes for Condition. The random-effects structure was then simplified following Bates, Kliegl, Vasishth, and Baayen (2015) to avoid overparametrization. The procedure used to simplify the random-effects structure of each model is described in the supplemental material, along with the final model outputs.

3.3.2 Results and discussion

		Experiment 1	Experiment 2	Experiment 3
Picture 1 accuracy (sd)	*Unrelated Condition*	82.00 (0.38)	81.38 (0.38)	79.18 (0.41)
	Long lag Condition	81.00 (0.38)	79.94 (0.39)	79.43 (0.40)
	Short lag Condition	81.56 (0.38)	81.50 (0.38)	80.63 (0.39)
	No lag Condition	84.06 (0.36)	80.94 (0.39)	-
Picture 1 latency (sd)	*Unrelated Condition*	717 (187)	860 (263)	854 (286)
	Long lag Condition	724 (191)	868 (251)	878 (323)
	Short lag Condition	715 (195)	847 (242)	850 (283)
	No lag Condition	707 (177)	870 (186)	-
Picture 2 accuracy (sd)	Unrelated Condition	70.25 (0.46)	67.44 (0.47)	67.48 (0.47)
	Long lag Condition	70.88 (0.45)	69.75 (0.46)	68.11 (0.46)
	Short lag Condition	72.88 (0.44)	71.44 (0.45)	71.26 (0.45)
	No lag Condition	81.38 (0.39)	77.19 (0.42)	-
Picture 2 latency (sd)	Unrelated Condition	769 (214)	821 (263)	815 (238)
	Long lag Condition	758 (209)	819 (251)	812 (235)
	Short lag Condition	744 (191)	789 (242)	798 (239)
	No lag Condition	618 (159)	638 (186)	-

Table 3.1: Mean accuracy and naming latencies, with standard deviations in parentheses, for Picture 1 and Picture 2 in Experiment 1, 2 and 3.

Table 3.1 shows the average accuracy rates and naming latencies for Pictures 1 and 2 per condition, along with the standard deviations. Recall that Picture 1 was always unrelated to the prime and that, consequently, the accuracy rates and latencies should not depend on the priming condition. By contrast, for Picture 2, repetition priming was expected for the response latencies and possibly the accuracy rates. These effects, if present, should decrease across lags.

As can be seen in the table, these predictions were borne out: Picture 1 accuracy rates and latencies were similar across priming conditions, whereas Picture 2 accuracy rates and naming latencies showed evidence for priming. The statistical analyses confirmed these impressions: The model with accuracy rates for Picture 1 naming included Condition as the independent variable, the random structure included only a by-participant intercept. Condition was not a significant predictor ($\chi^2(3) = 3.74$, $p = .29$). The model for log-transformed Picture 1 naming latencies included Condition as the independent variable; the random-effects structure included by-Picture 1 and by-participant intercepts and slopes for Condition. Including Condition as a predictor did not improve model fit ($\chi^2(3) = 2.35$, $p = .50$). Thus, as anticipated, responses to Picture 1 were not systematically affected by Condition.

By contrast, effects of Condition were found for the responses to Picture 2: the model for accuracy rates included Condition as the independent variable; and the random-effects structure included by-Picture 2 and by-participant intercepts. Including Condition improved model fit ($\chi^2(3) = 86.90$, $p < .002$). All contrasts were significant. In particular, the comparison between the unrelated condition and all other conditions together (first contrast) showed that people were more likely to name pictures incorrectly in the unrelated condition. In turn, items in the long lag condition were more likely to yield incorrect responses than those in the short lag and the no lag conditions together (second contrast). Responses in the short lag condition included more errors than in the no lag condition (third contrast). This means that participants were more likely to name a picture correctly if they had been presented with the prime at any time earlier during the experiment. The priming benefit was therefore strongest in the no lag condition and decreased at longer lags.

The model for the log-transformed Picture 2 naming latencies included Condition as the independent variable, and the random-effects structure included by-participant and by-Picture 2 slopes and intercepts. Condition was a significant predictor ($\chi^2(3) = 76.23, p < .002$). Paralleling the results seen for Picture 2 accuracy, response latencies for Picture 2 were shorter when participants had heard the primes earlier in the experiment than when this was not the case. More specifically, the first contrast showed that naming latencies in the unrelated condition were slower than in all other conditions; the second contrast showed that the long lag condition was slower than the short and no lag condition; the third contrast showed that the short lag condition was slower than the no lag condition. As pointed out, the contrasts used in the analysis do not allow us to compare the unrelated condition against each of the priming conditions. However, the mean Picture 2 latencies in the unrelated and long lag conditions were almost identical, suggesting that in the long lag condition repetition benefits might have not occurred. We return to this point in the General Discussion.

The goals of this experiment were to establish that a priming effect would be obtained from the primes embedded in carrier sentences and to determine the stability, or decay, of the effect across lags. The data showed that the effect was the strongest when prime and target were in the same trial and decreased with more intervening items. These results serve as a benchmark for the evaluation of priming effects in Experiments 2 and 3, where primes were presented while the participants prepared utterances.

3.4 Experiment 2

Experiment 1 showed repetition benefits when primes were embedded in sentences. The strongest repetition benefits were obtained in the no lag condition and priming benefits decreased at longer lags. The goal of Experiment 2 was to determine whether the same pattern would hold when the primes were processed during a picture naming task. While the participants of Experiment 1 first listened to the sentence including the prime and then named both Picture 1 and Picture 2, the participants of Experiment 2 saw Picture 1 while listening to the prime word and immediately named it. This forced them to plan a spoken response during the encoding of the prime. Determining whether priming occurred in this situation is important because studies of turn-taking suggest that in conversation comprehension and production processes often run in parallel. The division of attention between these processes may affect the way spoken utterances are processed and the impact of spoken primes on word production. If repetition priming did not occur under such circumstances, its value as a tool for supporting fast utterance planning in conversation would be limited. As explained earlier, repetition priming during word planning was also investigated by Jongman and Meyer (2017). However, in their study, primes were repeated twelve times during the experiment (six times as distractor and six times as target), were presented as single words, and immediately preceded the targets. Prime repetition, the (lack of) context and the distance between prime and target have been shown to affect repetition priming effects (e.g., McKone, 1998; Oliphant, 1983; Ostergaard, 1998). Unlike Jongman and Meyer (2017), in the present study the prime words and the targets (our Pictures 2) only occurred once; the primes were presented in carrier sentences, and Pictures 2 occurred in the same trial as the prime, or after a short or longer lag.

3.4.1 Method

Participants

Forty participants (8 male, mean age: 23, range: 19-28) took part in the study and received 8€ as compensation. Inclusion criteria were the same as in Experiment 1. Participants recruited in Experiment 1 were not eligible to take part in Experiment 2.

Apparatus, materials, design and data analysis

The same experimental setup, materials, and design were used as in Experiment 1. The data analysis was done in the same way as described above.

Procedure

The structure of Experiment 2 was the same as Experiment 1. The only difference was the timing of the presentation of Picture 1. While in Experiment 1 Picture 1 was presented after the sentence, in Experiment 2 it appeared at the onset of the prime word in the sentence (see Figure 3.1). Picture 1 stayed on the screen for 350ms and was followed by a fixation cross up to the presentation of Picture 2, which remained on screen for 2s. The gap between the onset of the sentence and the onset of Picture 2 was 4s, as in Experiment 1.

3.4.2 Results and discussion

Average accuracy rates, condition means and standard deviations for Picture 1 and 2 are shown in Table 3.1. As can be seen, the results are very similar to those obtained in Experiment 1. Again, the accuracy rates and naming latencies for Picture 1 were largely unaffected by Condition, whereas the accuracy rates and latencies for Picture 2 showed evidence for priming.

We first analysed Picture 1 accuracy rates. The model included Picture 1 Accuracy as the dependent variable and Condition as the independent variable. The random-effects structure included by-participant and by-Picture 1 slopes. The model with Condition as predictor did not improve over the null model ($\chi^2(3) = 2.50, p = .48$). Naming latencies in response to Picture 1 were also unaffected by Condition, as expected ($\chi^2(3) = 3.34, p = .34$). This model included log-transformed naming latencies as the dependent variable and Condition as the independent variable. The random-effects structure included by-participant and by-Picture 1 slopes and intercepts.

We then turned to the analysis of Picture 2 responses. We first modelled Accuracy rates. Condition was taken as the independent variable; the random-effects structure included a by-Picture 2 slope and by-participant and by-Picture 2 intercepts. Condition improved model fit ($\chi^2(3) = 27.17, p < .002$). All contrasts were significant. Similar results were obtained in the analysis of log-transformed naming latencies in response to Picture 2. Again, Condition was taken as the independent variable; the random-effects structure included a by-participant slope and by-participant and by-Picture 2 intercepts. Condition improved model fit

($\chi^2(3) = 73.96$, $p < .002$). As in the analysis of accuracy rates, all contrasts were significant. Again, this means that naming latencies in the unrelated condition were the slowest, that naming latencies in the long lag condition were slower than those in the short lag and no lag condition together, and that naming latencies in the short lag condition were slower than those in the no lag condition. Even if these results cannot give us specific information about the priming condition at which repetition priming can no longer be observed, mean naming latencies in the long lag condition were virtually identical to those in the unrelated condition, suggesting that at this point repetition did not yield any benefits, just as in Experiment 1. Overall though, our results suggest that, even when the cognitive load of the encoding task was increased by having participants perform the comprehension task together with a picture naming task, priming effects still occurred both in terms of higher accuracy rates and decreased latencies.

Finally, we compared the repetition benefits for the Picture 2 latencies in Experiment 1 and Experiment 2. The model included log-transformed naming latencies as the dependent variable, and Condition, Experiment and their interaction as the independent variables. The random-effects structure included by-participant slopes and intercepts for Condition and by-Picture 2 slopes and intercepts for Condition and Experiment. While both Condition ($\chi^2(3) = 145.02$, $p < .002$) and Experiment ($\chi^2(1) = 5.78$, $p = .02$) improved model fit, their interaction did not ($\chi^2(3) = 3.32$, $p = .35$). This means that repetition affected naming latencies equally in Experiments 1 and 2, in spite of the additional load imposed by the planning task in Experiment 2. Indeed, the priming effects for Picture 2 were very similar across experiments. The difference between the unrelated and the no lag condition was 151ms in Experiment 1 and 183ms in Experiment 2. The difference between the unrelated condition and the short lag condition was 26ms in Experiment and 31ms in Experiment 2, and as far as the long lag condition is concerned, the difference between the unrelated condition and the long lag condition was 11ms in Experiment 1 and 2ms in Experiment 2, suggesting that at this point repetition might not be beneficial for picture naming.

While the priming effects did not differ across experiments, the average naming latency for Picture 2 was longer in Experiment 2 than in Experiment 1. Since Picture 2 was always named in silence, without any concurrent task, we had not predicted a main effect of Experiment. The fact that latencies were longer in Experiment 2 than in Experiment 1 suggests that some degree of interference between speech-planning and comprehension still arose in Experiment 2 when

Picture 2 was named. We hypothesize that slower Picture 2 naming in Experiment 2 resulted from the additional processing load arising during the parallel (rather than sequential) processing of Picture 1 and the sentence, which then spilled over to Picture 2 naming latencies. Indeed, Picture 1 mean naming latencies in Experiment 2 were almost 150ms slower than in Experiment 1 (716ms in Experiment 1 and 862ms in Experiment 2), indicating that Picture 1 naming was indeed more efficient in silence than during sentence comprehension.

3.5 Experiment 3

Experiments 1 and 2 yielded robust repetition priming effects, regardless of whether the participants heard the prime and then named a picture or heard it while naming a picture. The comparison between the experiments revealed that dividing attention between comprehension and production did not affect the size of priming effects at any of the lags. In these experiments, the relationship between prime and Picture 2 was rather obvious because on no lag trials Picture 2 immediately followed the priming sentence. Participants might have been encouraged to pay attention to the spoken sentences because on these trials processing the prime was beneficial for the subsequent picture naming task. The goal of Experiment 3 was to measure priming effects when the relationship between prime and Picture 2 was less obvious. This was accomplished by removing trials where Picture 2 immediately followed the prime and only including the short and the long lag condition (lags of 10 and 50 trials).

3.5.1 Method

Participants

We recruited forty participants (10 male, mean age: 23.03 years, range: 18-28 years), who had not taken part in Experiments 1 or 2. Inclusion criteria and compensation were the same as in Experiments 1 and 2.

Apparatus, materials and design

Design, apparatus and scoring in Experiment 3 were the same as Experiments 1 and 2.

Procedure

Experiment 3 was identical to Experiment 2, with the only exception that the no lag condition was removed, which resulted in Condition having only three levels (short lag, long lag, unrelated).

Materials

The no lag condition trials and the 25 filler trials with identical primes and Picture 2 used in Experiments 1 and 2 were turned into unrelated filler trials by substituting the sentences containing the identical prime with new sentences containing unrelated prime words (mean frequency: 21.26, range: 1.35 – 128.70 per million in the SUBTLEX database, Keuleers et al. (2010)).

3.5.2 Results and discussion

Mean accuracy rates, naming latencies and standard deviations for Picture 1 and Picture 2 are reported in Table 3.1. Condition did not affect either Picture 1 accuracy rates or naming latencies. We first modelled Accuracy with respect to Condition. The random-effects structure included by-participant and by-Picture 1 intercepts. Condition did not improve model fit ($\chi^2(2) = 1.07$, $p = .59$). We then ran a model where the dependent variable was log-transformed Picture 1 latency and the independent variable was Condition. Again, the random-effects structure included by-participant and by-Picture 1 slopes and intercepts. Condition did not improve model fit ($\chi^2(2) = 2.67$, $p = .26$).

We then turned to the analyses of Picture 2 accuracy rates. The model included Picture 2 accuracy as the dependent variable and Condition as the independent variable. The random-effects structure included by-participant and by-Picture 2 intercepts. Condition improved model fit ($\chi^2(2) = 9.14$, $p = .01$). All contrasts were significant.

The same pattern was seen in the analysis of Picture 2 naming latencies. The model included log-transformed Picture 2 naming latencies as the dependent variable and Condition as the independent variable. The random-effects structure included by-participant and by-Picture 2 intercepts. Condition improved model fit ($\chi^2(2) = 13.00$, $p = .001$); all contrasts were significant, as in the case of Picture 2 accuracy. The first contrast compared the responses in the unrelated condition against responses in the long and short lag conditions together. Responses were slower in the unrelated condition than in the long and short lag conditions. The second contrast showed that naming latencies in the long lag conditions were slower than those in the short lag condition. We therefore concluded that the priming effects obtained in this study did not depend on the fact that participants strategically paid attention to the prime as a way to improve their performance in the immediately following picture naming task.

3.6 General Discussion

Mutual priming between interlocutors has been argued to contribute considerably to the smooth flow of everyday conversations (Garrod & Pickering, 2009; Pickering & Garrod, 2004): using words and structures that the partner has just produced can facilitate speech planning, both by guiding what to say and thereby contributing to building up common ground with the partner and by speeding up utterance planning. There is a large literature reporting robust repetition priming effects under laboratory conditions (e.g., Coane & Balota, 2010; Francis et al., 2008; Monsell et al., 1992; Wheeldon & Monsell, 1992). However, as these conditions differ in many ways from those prevailing in everyday conversation, we cannot take for granted that repetition priming supports speaking in conversation as effectively as the laboratory studies suggest. We cannot mimic spontaneous conversation in tightly controlled experiments. However, we can experimentally investigate variables that might affect the strength of repetition priming effects and impact its effectiveness in conversation. By doing so, we gain evidence about the mechanisms underlying repetition priming and its importance for conversation.

Following this logic, the current study investigated under which conditions repetition priming occurred from hearing word primes to naming target pictures. A closely related study by Jongman and Meyer (2017) had already established repetition priming from hearing single words to picture naming. Here we extended this work by presenting the prime words in sentence contexts and by varying the lag between prime and target, in terms of time and in terms of the amount of intervening materials. As in the earlier study, we examined whether priming occurred when participants merely listened to the primes and when they prepared to name distractor pictures while hearing the primes.

The results of the three experiments reported above are clear-cut: repetition priming facilitated picture naming. The priming effect was seen most strongly in the picture naming latencies but also emerged in the accuracy rates. The effect was moderated by the lag between prime and target, with priming being strongest at immediate repetition. Whether or not the participants prepared to name another picture while hearing the prime had little impact on the strength of the priming effect. In the remainder of this discussion, we compare these results to those of earlier studies and discuss the implications for our understanding of repetition priming and its potential role in conversation.

3.6.1 Spoken words prime picture naming

As described in the Introduction, in most studies of repetition priming in word production participants produced the primes as well as the targets (see Francis, 2014, for a review). Comprehending and producing words involve some shared or closely related representations, but the underlying processes are not identical. If the strength of repetition priming effects depends on the degree of overlap between prime and target representations and processing, heard primes should have weaker effects than self-produced primes onto picture naming. Moreover, merely listening to primes does not require the engagement of attention to the stimuli in the same way as naming primes does, which may also affect the strength of the priming effects. In sum, it is not evident that repetition effects as robust and strong as those seen when participants produce prime words will arise when prime words are merely heard.

The study by Jongman and Meyer (2017) already established that heard primes lead to robust repetition priming for picture naming. The present study replicates this finding. It is consistent with earlier findings by Wheeldon and Monsell (1992), who reported repetition priming for picture naming when the production of the primes was elicited using a definition task (e.g., "*Building in which horses are kept*" or "*An ___ a day keeps the doctor away*"). It is also consistent with the results obtained by Brown et al. (1991), who found comparable priming effects regardless of whether primes were overtly named or not. All of these findings indicate that repetition facilitates lexical access to the picture name (e.g., Barry et al., 2006; Francis et al., 2014; Monsell et al., 1992; Wheeldon & Monsell, 1992). However, they do not reveal which components of lexical access benefit most from repetition priming and through which mechanisms the facilitatory effect arises. These issues need to be assessed in further research.

As repetition priming effects are often relatively long-lived, they have been conceptualised as implicit learning processes involving lasting changes of the activation levels of lexical representations and the links between them (e.g., Hughes & Schnur, 2017; Monsell et al., 1992). As will be further discussed below, the repetition priming effects in the present study were short-lived, compared to those seen in other studies. This may suggest that changes in the representations of words induced by hearing them are short-lived or that part of the repetition priming effect seen here was due to the maintenance of explicit episodic representations of the primes, which decayed rapidly (e.g., K. Bock & Griffin, 2000).

3.6.2 Contextual embedding does not eliminate repetition priming from spoken words

Previous studies found that contextual embedding of primes reduced or even eliminated repetition priming effects (e.g., Coane and Balota (2010); Levy and Kirsner (1989); MacLeod (1989); Masson and Macleod (2000); Oliphant (1983); Speelman et al. (2002)). An account of this decrement of priming effects is that embedding of primes in contexts affects the distinctiveness of the memory traces. The earlier studies had not used the combination of spoken primes and pictorial targets used in the present study, but distinctiveness should be of relevance here as well. Therefore one might expect to see weaker repetition priming effects than in the earlier study by Jongman and Meyer (2017).

Contrary to this prediction, the repetition priming effect in the no lag condition, which was the only lag condition included in both studies, was numerically stronger in the present than in the earlier study (151ms and 181ms in Experiment 1 and 2 of the present study, 82ms in Experiment 1 of Jongman and Meyer, 2017). We refrain from any interpretation of this difference in the magnitude of the effects as the two studies differed in many respects. However, we can conclude that sentential embedding did not eliminate the repetition priming effect.

The absence of an effect of sentential context in the present study does not necessarily mean that the distinctiveness account is incorrect. Instead it indicates the need for further studies of the circumstances under which contextual embedding reduces the distinctiveness of individual words and their effectiveness as primes. In the current study the prime words were very prominent, always being the only noun of the sentence and the direct object of the verb. By contrast, earlier studies used more varied sentences featuring multiple nouns. In the present experiments the prime words may have "stuck out" more as particularly distinctive, which may have eliminated any contextual embedding effects. This account fits with the findings that, when embedded in text, low-frequency words yield more priming than high frequency words, and incongruent words, which do not fit in well with the context, yield more priming than congruent words (MacLeod, 1989; see also Coane Balota, 2010).

For the present purposes, it is most important to note that heard word primes, presented in isolation or in sentence contexts, led to substantial repetition priming effects for word production. This is consistent with the view that repetition priming can support speech production in conversation, regardless of whether the words produced by a conversational partner occur in isolation or in utterance contexts.

3.6.3 The effect of lag on contextually embedded spoken primes

In this study, we included two different lag conditions to compare the rate of decay of the priming effect in Experiment 1 and in Experiment 2. The type of contrasts that we used in our analysis does not allow us to clearly establish at which lag the priming effect disappeared. While we do not intend to make any substantial claims about the longevity of the effect, in this paragraph we will discuss our results in light of previous studies that used a similar paradigm and will outline some of the factors that may affect the longevity of the priming effect.

While mean naming latencies in the no lag and short lag condition were numerically smaller than those in the unrelated condition, mean naming latencies in the long lag condition were virtually identical to those in the unrelated condition. Although no formal comparison was carried out, this suggests that at that point the priming effect had probably faded. This pattern may be unexpected, given that some studies have reported priming effects lasting days or years (Cave, 1997; Mitchell, 2006). However, these studies used paradigms that differed substantially from the present one, most importantly perhaps in the use of prime pictures rather than words. The long-lived effect may therefore be based primarily on stable memory representations for the pictures rather than their names. Moreover, memory performance may have been improved by consolidation during sleep (Walker & Stickgold, 2004), which did not occur during the present study.

The effects of within-experiment lags between primes and targets have been assessed in a number of earlier studies. Consistent with the present results, they found decreasing priming effects with increasing lags. Nonetheless, in previous studies, repetition priming effects were observed at lags exceeding our long lags in both time and amount of intervening materials. For instance, Wheeldon and Monsell (1992) observed repetition after delays of 6 to 12 minutes (60 to 120 intervening trials); Monsell et al. (1992) observed repetition priming after 8 to 20 minutes.

There are many differences between the earlier studies and the present one that may account for any differences in the longevity of the repetition priming effect. One difference is the presence or absence of the sentential embedding of the primes. Presentation of the primes in context does not eliminate the priming effect at short lags, but it may affect the rate of decay. This suggestion could be assessed in a study directly contrasting the effects of isolated word primes versus embedded primes at different lags. Another potentially important difference to the earlier studies is that our participants heard the primes, whereas the participants of the earlier studies produced the prime words themselves. Producing words at study leads to better memory performance than hearing or reading them (e.g., Fawcett, Quinlan, & Taylor, 2012; MacLeod & Bodner, 2017; MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010; Zormpa, Brehm, Hoedemaker, & Meyer, 2019). This holds for explicit memory tasks (free recall and recognition) and implicit memory tasks (fragment completion and picture naming; but see Kahn and Arnold (2015) for equal production-to-production and comprehension-to-production priming). To assess this proposal a study would have to contrast the effects of self-produced versus heard primes at different lags.

It is difficult to estimate on the basis of our findings how long repetition priming may last in everyday conversation. Turns in conversations vary greatly in length. The average duration of turns in casual conversation has been estimated to be about 2 seconds (one or two sentences; Levinson (2016)). We find repetition priming with a lag of one minute, corresponding to ten intervening trials. Thus a conservative estimate of the impact of repetition priming in conversation is that speakers should benefit from it at least during three or four turns following the prime. The impact of the prime may vary given the elapsed time as well as the amount and type of intervening items. Further research teasing apart the impact of these variables would be very valuable not only for a better understanding of the role of repetition priming in conversation, but also for elucidating the mechanisms underlying repetition priming.

3.6.4 Linguistic dual tasking does not affect repetition priming

Finally, we replicated the earlier finding that repetition priming occurs when participants plan words while listening to the primes. In fact, the joint analyses of the results of Experiments 1 and 2 showed that the priming effects in the two experiments did not differ significantly in strength. This replicates the findings by Jongman and Meyer (2017) and extends them by demonstrating that heard words also prime picture naming in a linguistic dual-tasking setting when the primes are embedded in sentence contexts and when there is a lag between prime and target. Work using different paradigms has also shown that priming occurs under conditions of divided attention, though the priming effect can be weaker than under full attention conditions (for a meta-analysis see Spataro et al. (2011)). The finding that the size of the repetition priming effect was largely unaffected by the division of attention is consistent with the suggestion that strategic processing of the primes did not contribute much to the priming effect.

The occurrence of repetition priming during speech planning should be followed up in further research. In the present study primes were processed sufficiently to generate priming effects, but, as their on-line processing was not tracked, we cannot say whether and how it was affected by the speech planning task. Prime processing might have been delayed or shallower. The study by Jongman and Meyer (2017) included an associative priming condition in addition to the repetition priming condition, and found that in one of the two experiments the associative priming effect was eliminated under linguistic dual-tasking. This suggests that the processing of the primes was affected either because attention had to be divided between comprehension and speech planning, or because of interference between representations accessed for production and comprehension. Furthermore, the distractor pictures were named more slowly in our Experiment 2 than in Experiment 1, and, correspondingly, in Jongman and Meyer's plan condition than in their no plan conditions. This also indicates that the processing of the spoken input was hindered by concurrent speech planning. Further investigating such interference effects would be important to gain a better understanding of the relationships between the speech comprehension and production system and the way they can be simultaneously engaged in conversation.

While the results of this experiment suggest that repetition priming can occur during linguistic dual-tasking, our experimental set-up differed in many ways from real-life conversational settings. One important difference was that in our experiments, the Picture 2-naming task was not contingent on the content of the spoken prime, whereas in in everyday conversations people typically respond to the content of their partner's utterances. This means that during conversations participants might prioritise comprehension processes more, which could strengthen repetition priming from comprehension to production and, therefore, lead to more alignment between speakers. However, in our study (especially in Experiments 1 and 2), participants may very well have been aware of the fact that many of the prime words were target names, which may have encouraged them to listen carefully to the primes. An important question for further research is how people distribute their attention in conversation, and how this affects mutual priming.

Another important difference to everyday conversation is that in our study, participants were required to produce the names of the target pictures using nouns that had or had not occurred before. By contrast, in conversation speakers often choose not to use full noun phrases to refer to concepts introduced before, but instead use pronouns (Arnold, 2010; Arnold & Zerkle, 2019). In fact, comprehension studies have shown a repeated name penalty, i.e., longer reading times for noun phrases in contexts where pronouns were expected (e.g., Kennison & Gordon, 1997). There is a substantial literature on the linguistic and cognitive variables that affect preferences for nouns or pronouns (e.g., Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000; Brown-Schmidt, Byron, & Tanenhaus, 2005; Fukumura & van Gompel, 2012). However, little is known about the cognitive processes occurring when speakers decide whether to use a noun or pronoun. Our data do not speak to this issue, but only show that, when a noun is to be produced, the speed of producing it can be reduced through repetition priming. An important question for further research is how speakers' choices between nouns and pronouns might be affected by linguistic dual-tasking.

3.7 Conclusions

Mutual repetition priming between interlocutors might support fluent conversation in at least two ways, i.e., by contributing to the establishment of common ground and by increasing the speed of speech planning. This pivotal role of priming can only be confirmed if it occurs under the conditions prevailing in conversation. We showed that repetition priming for word production occurred when primes were embedded in sentences, when the target word followed the prime immediately or after a short lag, and regardless of whether or not participants were planning to name a distractor picture while listening to the primes. These results suggest that repetition priming may indeed aid speech planning in conversation by reducing processing costs.

3.8 Supplemental materials

3.8.1 Procedure for random-effects justification

In each analysis, we first ran the most complex model to converge (as shown in the header of each table). PCA of the random-effects covariance matrix was then used to determine how much variance was explained by each random term (Bates et al., 2015). [,1] refers to the variance explained by each random intercept for Condition, [,2], [3,], [,4] to that explained by the slope for Condition (one for each level of the fixed predictor), and [,5] to the variance explained by the random intercept for Experiment. In each step described below, we first removed the random term that explain the smaller proportion of variance. Whenever the proportion of variance explained by each random term was > 0, we checked that removing the random term did not affect model fit by running a Likelihood Ratio Test comparing the model without and with the random term. Here, we also report the output of the selected final model for each analysis.

3.8.2 Experiment 1

Picture 1 accuracy

P1Accuracy ~ Condition + (Condition|Participant) + (Condition|Picture 1)					
		[,1]	[,2]	[,3]	[,4]
Cumulative proportion	Picture 1	.80	1.00	1.00	1.00
	Participant	1.00	1.00	1.00	1.00

Table 3.2: Variance explained by random-effects in the model with Picture 1 accuracy (correct/non correct) as the dependent variable and Condition as the independent variable (Experiment 1).

- We removed the by-participant slope ($\chi^2(9) = 3.61$, $p = .94$).
- We did not remove the by-Picture 1 slope ($\chi^2(9) = 26.43$, $p = .002$).
- This model (M1) performed better than the null model ($\chi^2(9) = 9.02$, $p = .03$). However, several iterations failed during confidence interval bootstrapping (>5%), we removed the by-Picture 1 slope. This model did not perform better than the null model ($\chi^2(3) = 6.62$, $p = .09$). Again, at least 5% of iterations failed during bootstrapping. The final model (M2) is the following: Accuracy ~ Condition + (1|Participant). In this case, 4.5% of iterations failed during bootstrapping of confidence intervals. This model did not perform better than the null model ($\chi^2(3) = 3.74$, $p =$

.29). We retained this model. Further confirmation of the fact that Picture 1 accuracy rates were independent of Condition comes from the results obtained in Experiment 2. Using the same materials, there was no effect of Condition on Picture 1 accuracy rates or latencies. Below we report the outputs of both M1 and M2.

Fixed effects					**Random effects**			
	Estimate	St. err.	z-value	CI			Variance	St. dev.
Intercept	2.38	0.16	15.05	2.06, 2.73	*Picture 1*	**Intercept**	2.44	1.56
						Condition 1	0.30	0.54
						Condition 2	0.36	0.60
						Condition 3	0.04	0.19
Condition 1	0.02	0.14	0.16	-0.27, 0.32	*Participant*	**Intercept**	0.17	0.42
Condition 2	-0.31	0.14	-2.30	-0.62, -0.006				
Condition 3	-0.36	0.16	-2.30	-0.72, -0.03				

Table 3.3: Output of the model (M1) with Picture 1 accuracy (correct/non correct) as the dependent variable and Condition as the independent variable (Experiment 1).

Fixed effects					**Random effects**			
	Estimate	St. err.	z-value	CI			Variance	St. dev.
Intercept	1.62	0.06	28.44	1.50, 1.73	*Participant*	**Intercept**	0.08	0.29
Condition 1	-0.06	0.08	-0.66	-0.20, 0.10				
Condition 2	-0.09	0.08	-1.03	-0.24, 0.07				
Condition 3	-0.15	0.10	-1.53	-0.34, 0.05				

Table 3.4: Output of the model (M2) with Picture 1 accuracy (correct/non correct) as the dependent variable and Condition as the independent variable (Experiment 1).

Picture 1 latencies

P1logOnset ~ Condition + (Condition|Participant) + (Condition|Picture 1)					
		[,1]	[,2]	[,3]	[,4]
Cumulative proportion	Picture 1	.57	.79	.91	1.00
	Participant	.74	.92	.99	1.00

Table 3.5: Variance explained by random-effects in the model with Picture 1 log-transformed naming latencies as the dependent variable and Condition as the independent variable (Experiment 1).

- We did not remove the by-participant slope ($\chi^2(9) = 34.60$, $p < .002$).
- We did not remove the by-picture 1 slope ($\chi^2(9) = 73.25$, $p < .002$).

Fixed effects				Random effects			
	Estimate	St. err.	t-value			Variance	St. dev.
Intercept	6.56	0.02	324.17	*Picture 1*	**Intercept**	0.02	0.13
					Condition 1	0.004	0.06
					Condition 2	0.003	0.06
					Condition 3	0.006	0.08
Condition 1	-0.003	0.01	-0.28	*Participant*	**Intercept**	0.01	0.11
					Condition 1	0.001	0.04
					Condition 2	0.002	0.04
					Condition 3	0.002	0.04
Condition 2	0.01	0.01	1.16				
Condition 3	0.01	0.01	1.12				

Table 3.6: Output of the model with Picture 1 log-transformed naming latencies as the dependent variable and Condition as the independent variable (Experiment 1).

Picture 2 accuracy

P2Accuracy ~ Condition + (Condition|Participant) + (1|Picture 2)					
		[,1]	[,2]	[,3]	[,4]
Cumulative proportion	Picture 2	1.00	-	-	-
	Participant	.72	1.00	1.00	1.00

Table 3.7: Variance explained by random-effects in the model with Picture 2 accuracy (correct/non correct) as the dependent variable and Condition as the independent variable (Experiment 1).

- We removed the by-participant slope ($\chi^2(9) = 5.80$, $p = .76$)

Fixed effects					Random effects			
	Estimate	St. err.	z-value	CI			Variance	St. dev.
Intercept	1.35	0.11	11.84	1.14, 1.59	*Picture 2*	**Intercept**	1.27	1.13
Condition 1	-0.33	0.07	-4.55	-0.48, -0.19	*Participant*	**Intercept**	0.15	0.39
Condition 2	-0.43	0.08	-5.57	-0.59, -0.28				
Condition 3	-0.62	0.10	-6.45	-0.81, -0.43				

Table 3.8: Output of the model with Picture 2 accuracy (correct/non correct) as the dependent variable and Condition as the independent variable (Experiment 1).

Picture 2 latencies

P2logOnset ∼ Condition + (Condition|Participant)+(Condition|Picture 2)					
		[,1]	[,2]	[,3]	[,4]
Cumulative proportion	Picture 2	0.69	0.93	0.98	1.00
	Participant	0.50	0.96	1.00	1.00

Table 3.9: Variance explained by random-effects in the model with Picture 2 log-transformed naming latencies as the dependent variable and Condition as the independent variable (Experiment 1).

- We did not remove the by-participant slope because it improved model fit ($\chi^2(9) = 108.25$, $p < .002$).
- We did not remove the by-participant slope because it improved model fit ($\chi^2(9) = 86.22$, $p < .002$).

Fixed effects				**Random effects**			
	Estimate	St. err.	t-value			Variance	St. dev.
Intercept	6.56	0.02	345.47	*Picture 2*	**Intercept**	0.02	0.12
					Condition 1	0.002	0.04
					Condition 2	0.005	0.07
					Condition 3	0.004	0.07
Condition 1	0.10	0.01	10.26	*Participant*	**Intercept**	0.01	0.10
					Condition 1	0.001	0.03
					Condition 2	0.004	0.06
					Condition 3	0.005	0.07
Condition 2	0.12	0.01	8.75				
Condition 3	0.20	0.02	13.39				

Table 3.10: Output of the model with Picture 2 log-transformed naming latencies as the dependent variable and Condition as the independent variable (Experiment 1).

3.8.3 Experiment 2

Picture 1 accuracy

P1Accuracy ∼ Condition + (Condition|Participant)+(1|Picture 1)					
		[,1]	[,2]	[,3]	[,4]
Cumulative proportion	Picture 1	1.00	-	-	-
	Participant	0.86	1.00	1.00	1.00

Table 3.11: Variance explained by random-effects in the model with Picture 1 accuracy (correct/non correct) as the dependent variable and Condition as the independent variable (Experiment 2).

- We removed the by-participant slope ($\chi^2(9) = 2.73$, p = .97).

Fixed effects					**Random effects**			
	Estimate	St. err.	z-value	CI			Variance	St. dev.
Intercept	2.08	0.14	14.69	1.80, 2.36	*Picture 1*	**Intercept**	1.68	1.30
Condition 1	0.10	0.09	1.19	-0.07, 0.28	*Participant*	**Intercept**	0.23	0.48
Condition 2	-0.10	0.09	-1.10	-0.27, 0.10				
Condition 3	0.01	0.10	0.08	-0.20, 0.21				

Table 3.12: Output of the model with Picture 1 accuracy (correct/non correct) as the dependent variable and Condition as the independent variable (Experiment 2).

Picture 1 latencies

P1logOnset ∼ Condition + (Condition|Participant)+(Condition|Picture 1)					
		[,1]	[,2]	[,3]	[,4]
Cumulative proportion	Picture 1	.39	.74	.96	1.00
	Participant	.78	.93	.98	1.00

Table 3.13: Variance explained by random-effects in the model with Picture 1 log-transformed naming latencies as the dependent variable and Condition as the independent variable (Experiment 2).

- We did not remove the by-participant slope ($\chi^2(9) = 41.02$, $p < .002$).
- We did not remove the by-Picture 2 slope because it increased model fit ($\chi^2(9) = 85.12$, $p < .002$).

Fixed effects				Random effects			
	Estimate	St. err.	t-value			Variance	St. dev.
Intercept	6.72	0.03	283.30	*Picture 1*	**Intercept**	0.01	0.11
					Condition 1	0.008	0.09
					Condition 2	0.003	0.05
					Condition 3	0.01	0.11
Condition 1	-0.005	0.01	-0.38	*Participant*	**Intercept**	0.03	0.17
					Condition 1	0.003	0.06
					Condition 2	0.003	0.06
					Condition 3	0.002	0.04
Condition 2	0.005	0.01	0.39				
Condition 3	-0.02	0.01	-1.61				

Table 3.14: Variance explained by random-effects in the model with Picture 1 log-transformed naming latencies as the dependent variable and Condition as the independent variable (Experiment 2).

Picture 2 accuracy

P2Accuracy ~ Condition + (Condition|Participant)+(Condition|Picture 2)					
		[,1]	[,2]	[,3]	[,4]
Cumulative proportion	Picture 2	.42	.71	.90	1.00
	Participant	.99	1.00	1.00	1.00

Table 3.15: Variance explained by random-effects in the model with Picture 2 accuracy (correct/non correct) as the dependent variable and Condition as the independent variable (Experiment 2).

- We removed the by-participant slope ($\chi^2(9) = 2.32$, $p = .99$).
- We did not remove the by-Picture 2 slope ($\chi^2(9) = 51.89$, $p < .002$).

Fixed effects					Random effects			
	Estimate	St. err.	z-value	CI			Variance	St. dev.
Intercept	1.24	0.12	10.06	0.99, 1.47	*Picture 2*	**Intercept**	1.15	1.07
						Condition 1	0.51	0.71
						Condition 2	0.38	0.61
						Condition 3	0.74	0.86
Condition 1	-0.31	0.10	-3.25	-0.52, -0.11	*Participant*	**Intercept**	0.27	0.52
Condition 2	-0.32	0.10	-3.24	-0.50, -0.12				
Condition 3	-0.38	0.13	-3.04	-0.64, -0.13				

Table 3.16: Output of the model with Picture 2 accuracy (correct/non correct) as the dependent variable and Condition as the independent variable (Experiment 2).

Picture 2 latencies

P2logOnset ∼ Condition + (Condition|Participant)+(Condition|Picture 2)					
		[,1]	[,2]	[,3]	[,4]
Cumulative proportion	Picture 2	.80	.94	1.00	1.00
	Participant	.52	.98	1.00	1.00

Table 3.17: Variance explained by random-effects in the model with Picture 2 log-transformed naming latencies as the dependent variable and Condition as the independent variable (Experiment 2).

- We did not remove the by-Picture 2 slope ($\chi^2(9) = 76.78$, $p < .002$).
- We did not remove the by-participant slope ($\chi^2(9) = 98.89$, $p < .002$).
- Since the null model with the full random-effects structure did not converge, we removed the by-Picture 2 slope.

Fixed effects				**Random effects**			
	Estimate	St. err.	t-value			Variance	St. dev.
Intercept	6.61	0.02	324.33	*Picture 2*	**Intercept**	0.02	0.14
Condition 1	0.10	0.01	10.42	*Participant*	**Intercept**	0.01	0.11
					Condition 1	0.002	0.04
					Condition 2	0.003	0.06
					Condition 3	0.008	0.09
Condition 2	0.14	0.01	11.74				
Condition 3	0.22	0.02	12.61				

Table 3.18: Output of the model with Picture 2 log-transformed naming latencies as the dependent variable and Condition as the independent variable (Experiment 2).

Comparison between Experiment 1 and 2 latencies

P2logOnset ∼ Condition*Experiment + (1+Condition |Participant)+(1+Condition+Experiment|Picture 2)						
		[,1]	[,2]	[,3]	[,4]	[,5]
Cumulative proportion	Picture 2	.75	.92	.99	1.00	1.00
	Participant	.51	1.00	1.00	1.00	-

Table 3.19: Variance explained by random-effects in the model with Picture 2 log-transformed naming latencies as the dependent variable and Condition, Experiment and their interaction as the independent variable (Experiments 1 and 2)

- We did not remove the by-Picture 2 slope for Experiment ($\chi^2(5) = 22.63$, $p < .002$).

- We did not remove the by-Picture 2 slope for Condition ($\chi^2(12) = 182.03$, $p < .002$).
- We did not remove the by-Participant slope for Condition ($\chi^2(9) = 212.7$, $p < .002$).

Fixed effects				**Random effects**			
	Estimate	St. err.	t-value			Variance	St. dev.
Intercept	6.59	0.02	417.44	*Picture 2*	**Intercept**	0.02	0.13
Condition 1	0.10	0.01	12.88		**Condition 1**	0.004	0.06
Condition 2	0.13	0.01	13.89		**Condition 2**	0.003	0.06
Condition 3	0.21	0.01	17.05		**Condition 3**	0.01	0.07
					Experiment	0.0001	0.01
Experiment 1	-0.03	0.01	-2.48	*Participant*	**Intercept**	0.01	0.10
					Condition 1	0.001	0.04
					Condition 2	0.004	0.06
					Condition 3	0.01	0.08

Table 3.20: Output of the model with Picture 2 log-transformed naming latencies as the dependent variable and Condition and Experiment as the independent variables (Experiments 1 and 2).

3.8.4 Experiment 3

Picture 1 accuracy

P1Accuracy ∼ Condition + (Condition|Participant)+(Condition|Picture 1)				
		[,1]	[,2]	[,3]
Cumulative proportion	Picture 1	1.00	1.00	1.00
	Participant	.76	1.00	1.00

Table 3.21: Variance explained by random-effects in the model with Picture 1 accuracy (correct/non correct) as the dependent variable and Condition as the independent variable (Experiment 3).

- We removed the by-Picture 1 slope ($\chi^2(5) = 1.06$, $p = .96$)
- We removed the by-participant slope ($\chi^2(5) = 1.51$, $p = .91$)

Fixed effects					**Random effects**			
	Estimate	St. err.	z-value	CI			Variance	St. dev.
Intercept	2.06	0.16	12.59	1.73, 2.39	*Picture 1*	**Intercept**	2.25	1.50
Condition 1	-0.06	0.09	-0.70	-0.24, 0.11	*Participant*	**Intercept**	0.30	0.55
Condition 2	-0.08	0.11	-0.80	-0.32, 0.11				

Table 3.22: Output of the model with Picture 1 accuracy (correct/non correct) as the dependent variable and Condition as the independent variable (Experiment 3).

Picture 1 latencies

P1logOnset ∼ Condition + (Condition|Participant)+(Condition|Picture 1)				
		[,1]	[,2]	[,3]
Cumulative proportion	Picture 1	.58	.88	1.00
	Participant	.75	.95	1.00

Table 3.23: Variance explained by random-effects in the model with Picture 1 log-transformed naming latencies as the dependent variable and Condition as the independent variable (Experiment 3).

- We did not remove the by-participant slope because it increased model fit ($\chi^2(5) = 40.90$, $p < .002$).
- We did not remove the by-Picture 1 slope because it increased model fit ($\chi^2(5) = 35.79$, $p < .002$).

Fixed effects				**Random effects**			
	Estimate	St. err.	t-value			Variance	St. dev.
Intercept	6.72	0.03	243.85	*Picture 1*	**Intercept**	0.02	0.12
					Condition 1	0.007	0.08
					Condition 2	0.005	0.07
Condition 1	-0.004	0.02	-0.28	*Participant*	**Intercept**	0.03	0.16
					Condition 1	0.005	0.07
					Condition 2	0.005	0.07
Condition 2	0.03	0.02	1.62				

Table 3.24: Output of the model with Picture 1 log-transformed naming latencies as the dependent variable and Condition as the independent variables (Experiment 3).

Picture 2 accuracy

P2Accuracy ~ Condition + (Condition|Participant)+(Condition|Picture 2)				
		[,1]	[,2]	[,3]
Cumulative proportion	Picture 2	.84	.95	1.00
	Participant	1.00	1.00	1.00

Table 3.25: Variance explained by random-effects in the model with Picture 2 accuracy (correct/non correct) as the dependent variable and Condition as the independent variable (Experiment 3).

- We removed the by-participant slope ($\chi^2(5) = 1.39$, $p = .92$).
- We removed the by-Picture 2 slope ($\chi^2(5) = 4.87$, $p = .43$).

Fixed effects					**Random effects**			
	Estimate	St. err.	z-value	CI			Variance	St. dev.
Intercept	1.07	0.14	7.77	.79, 1.35	*Picture 2*	**Intercept**	1.50	1.22
Condition 1	-0.17	0.08	-2.22	-.33, -.02	*Participant*	**Intercept**	0.32	0.56
Condition 2	-0.19	0.09	-2.13	-.36, -.003				

Table 3.26: Output of the model with Picture 2 accuracy (correct/non correct) as the dependent variable and Condition as the independent variable (Experiment 3).

Picture 2 latencies

P2logOnset ~ Condition + (Condition|Participant)+(Condition|Picture 2)				
		[,1]	[,2]	[,3]
Cumulative proportion	Picture 2	.84	.93	1.00
	Participant	.85	.96	1.00

Table 3.27: Variance explained by random-effects in the model with Picture 2 log-transformed naming latencies as the dependent variable and Condition as the independent variable (Experiment 3).

- We removed the by-participant slope ($\chi^2(5) = 9.50$, $p = .09$).
- We removed the by-Picture 2 slope ($\chi^2(5) = 10.87$, $p = .05$).

Fixed effects				**Random effects**			
	Estimate	St. err.	t-value			Variance	St. dev.
Intercept	6.67	0.02	316.55	*Picture 2*	**Intercept**	0.02	0.16
Condition 1	0.02	0.01	2.23	*Participant*	**Intercept**	0.01	0.11
Condition 2	0.02	0.01	2.81				

Table 3.28: Output of the model with Picture 2 log-transformed naming latencies as the dependent variable and Condition as the independent variables (Experiment 3).

4 The effect of linguistic dual-tasking on processing spoken words: evidence from EEG

Abstract

While previous studies found that concurrent comprehension affects production processes, less is known about any effects of production on comprehension. Using EEG, we investigated whether concurrent production impacts on online comprehension processes and repetition priming. The experiment included a study phase, during which participants heard words passively (NP block, block 1) or while planning an unrelated response (P block, block 2), and a test phase, during which participants named pictures whose name could be identical or unrelated to the words used in the study phase. In the test phase, we only found very weak repetition priming effects, which did not differ across blocks. In the study phase, we found a reduced N1 in the NP block compared to the P block, which we interpreted as an effect of task order. We also found a positive cluster, which we interpreted as an increased positivity in the P block compared to the NP block. We discuss the theoretical and methodological implications of the findings.

Keywords: speech planning, linguistic dual-tasking, repetition priming, N400, positivity

4.1 Introduction

Dual-tasking is ubiquitous in our daily lives: for instance, we listen to lectures while taking notes, we walk while checking our phones, or we eat while watching tv. Conversation is also a type of dual-tasking. Indeed, there is evidence that speakers can prepare a response while their interlocutor has not completed their turn yet (Barthel et al., 2016; Bögels et al., 2018; Corps et al., 2018; Levinson & Torreira, 2015; Sjerps & Meyer, 2015).

Dual-tasking is not as efficient as single-tasking (Pashler, 2000; Worringer et al., 2019). For instance, speaking or listening while carrying out a secondary task often results in performance decrements on the secondary task or both (e.g., Becic et al., 2010; Boiteau et al., 2014; Ferreira, Bailey, & Ferraro, 2002; Strayer & Johnston, 2001). In order to shed light on how production and comprehension unfold in conversations, a few psycholinguistic studies have investigated whether linguistic dual-tasking (i.e., simultaneous speech planning and comprehension) impacts on comprehension and/or production processes. The deleterious effect of linguistic dual-tasking on *production* is quite clear. A number of studies have shown that participants are slower at producing a spoken response when they are listening to speech at the same time than when they are planning the response in silence (e.g., Bögels et al., 2015; Fairs et al., 2018; Fargier & Laganaro, 2016). By contrast, studies on the effect of linguistic dual-tasking on *comprehension* did not find conclusive evidence that comprehension processes are affected by concurrent speech planning (Bartolozzi et al., 2021; Bögels et al., 2018; Daliri & Max, 2016; Fargier & Laganaro, 2019; Jongman & Meyer, 2017; Martin, Branzi, & Bar, 2018).

We investigated whether online comprehension processes (indexed by event-related brain responses) are affected by a concurrent speech planning task. A few experiments have suggested that this may be the case (Daliri & Max, 2016; Fargier & Laganaro, 2019). In addition, we investigated whether linguistic dual-tasking can affect how easily words are later re-accessed and used. More specifically, we used a repetition priming paradigm to investigate whether linguistic dual-tasking during the study phase affected the magnitude of repetition priming at test. We measured repetition priming effects for words encoded during simple comprehension or linguistic dual-tasking and determined whether the encoding condition affected the magnitude of repetition priming effects. Before turning to the details of the study, we first describe previous work that investigated the effect of concurrent speech planning on both online comprehension processes and priming.

4.1.1 Effect of linguistic dual-tasking on online comprehension processes

Linguistic dual-tasking can affect comprehension in two ways: first of all, it can influence online comprehension processes occurring at word presentation. In studies on linguistic dual-tasking, these online processes can be described using EEG measures. In addition to online comprehension processes, linguistic dual-tasking can also affect offline processes, which we define as the access and retrieval of words initially comprehended during linguistic dual-tasking (e.g., priming effects or episodic retrieval). As far as we are aware, previous studies that looked at the effect of linguistic dual-tasking on online comprehension either measured N1 or N400 differences between single-tasking and dual-tasking paradigms. The auditory N1 is an early negative deflection peaking around 100ms post stimulus onset and it is the result of three different constituents, with the biggest contribution at fronto-central sites (Luck & Kappenman, 2011). The N100 has been linked to early perception of auditory information and it can be affected by attentional modulations (Joos, Gilles, Van de Heyning, De Ridder, & Vanneste, 2014). The N400 is a negative-going component usually peaking 400ms after stimulus onset and with a centro-parietal distribution (Kutas & Federmeier, 2011). The N400 is usually associated with processing of meaningful stimuli (Kutas & Federmeier, 2011).

Studies investigating the effects of linguistic dual-tasking on the N1 exclusively used tones and/or syllables in the comprehension task. A first study by Daliri and Max (2016), who recorded EEG during a delayed naming task, showed that early auditory comprehension processes are hindered by concurrent phonological and/or articulatory processes involved in production. Participants were presented with words, which changed colour after a delay. In one condition, participants read the words in silence, in another condition they read them aloud when they changed colour. In some of the trials, a tone or a syllable (spoken by the same participant at the beginning of the session and recorded) was played during the delay. Long latency auditory evoked potentials were recorded in this delay period. The amplitude of the auditory N100 an P200 in response to the syllables was smaller when participants were planning to read the word aloud rather than in silence. By contrast, the tones were only accompanied by a modulation of the N100 but not of the P200. Overall, the authors suggested that the modulation of the N100 for both syllables and tones and of the P200 for syllables may be due to the central nervous system trying to prepare the auditory system for auditory feedback during overt word reading. Furthermore, the fact that tones elicited only N100 modulations, while syllables elicited both N100 and P200 modulations, suggests that some information at the syllable level is affected by the production task.

In another study using syllables as auditory stimuli, Fargier and Laganaro (2019) compared neural correlates of picture-naming and concurrent syllable processing with respect to single-tasking. Participants named pictures while simultaneously listening to syllables played 150ms, 300ms or 450ms post picture onset. Participants were asked to either hear syllables passively (passive hearing) or to monitor the occurrence of a specific syllable (active listening). In addition to these two dual-tasking conditions, the experiment included two single-tasking conditions in which participants named pictures, and heard syllables without a concurrent production task. While the P2 component did not differ across conditions, the N1 had a longer latency and a greater amplitude when participants heard syllables while naming pictures than when they heard syllables in silence. However, this was only evident at the longest SOA, 450ms, that is to say when syllabification or phonetic encoding in the picture naming task (Indefrey, 2011; Indefrey & Levelt, 2004) was carried out together with early comprehension processes. Moreover, during picture-naming while hearing syllables at SOA=450ms, fast naming latencies in the picture-naming task were associated with longer latencies of the global field power (GFP, for a definition, see Murray, Brunet, and Michel (2008)). The naming latencies in the picture-naming task reflected production processes, while the global field power described syllable processing and, in general, comprehension processes. The trade-off between production and comprehension therefore suggests interference of production on comprehension processes. The finding that interference of speech planning on comprehension was only evident at the longest SOA may suggest that the effect is not only driven by division of attention between tasks but also by the type of production and comprehension subprocesses carried out at the same time.

While these studies looked at the effect of production on early auditory processing (as indexed by N1 modulations), they could not speak to whether processing of meaningful stimuli in comprehension may also be disrupted during concurrent speech planning. This question was partly addressed by Martin et al. (2018) and Bögels et al. (2018), who used a dual-tasking paradigm where participants listened to sentence while producing either a syllable (Martin et al., 2018) or a meaningful response (Bögels et al., 2018). The aim of these two studies was to determine whether prediction processes during comprehension, indexed by an N400 effect for unpredictable vs predictable words, were affected by linguistic dual-tasking.

Martin et al. (2018) asked participants to read Spanish sentences ending with unexpected or expected words. Unexpected and expected words had different grammatical gender (e.g., *una corona*, a crown, vs *un sombrero*, a hat). As a result, predictability effects should already be evident at the onset of the article preceding the target word. According to the group they were assigned to, participants had to read the sentences while 1) repeating a syllable aloud multiple times, 2) listening to a syllable being repeated multiple times, or 3) clicking their tongue multiple times. Prediction processes during sentence reading, indexed by an increased N400 for unpredictable vs non-predictable endings (measured at the article onset), were disrupted when participants concurrently repeated the syllable, but not when they listened to a repeated syllable or when they clicked their tongue. The authors explained that prediction processes make use of speech production resources and that, when the latter are taxed, prediction is hindered.

Using a more naturalistic paradigm, Bögels et al. (2018) had participants perform a joint task with a confederate: the confederate asked the participant a question (e.g., *Welk object is krom en wordt als fruit gezien?*, Which object is curved and is considered a type of fruit?). The answer indicated which item out of those presented on the screen (e.g., a banana and a pineapple) the confederate had to select. The answer to the question could be planned early or late, according to position of the informative cue in the question. Each question also contained an expected or an unexpected word (*fruit* vs *gezond* (healthy) in the above example), which could occur early or late in the question (before or after planning could start). The authors argued that, if comprehension is less efficient during concurrent speech planning, the N400 for unexpected words should be greater when the unexpected word occurs before (rather than after) participants can start planning (Kutas & Federmeier, 2011). While this prediction was not borne out, the N400 effect in the early plan condition (but not in the late plan condition) correlated with the mean response time (across conditions) of each participant. In the early plan condition, the unexpected or expected word occurred after the critical information necessary to answer the question, i.e., after planning could start. The correlation between the N400 effect and mean response time suggested that, in the early plan condition, participants who prioritised speech planning over comprehension poorly attended to the remaining part of the question, which resulted in a reduced N400 effect for unexpected vs expected words.

In sum, studies of the effect of linguistic dual-tasking on online comprehension point at two main conclusions. First, early auditory comprehension processes are hindered by concurrent post-lexical processing in speech production (Daliri & Max, 2016; Fargier & Laganaro, 2019). Second, prediction processes may also be disrupted during linguistic dual-tasking (Bögels et al., 2018; Martin et al., 2018). While any difference between these two studies may be due to discrepancies between the production tasks used (continuous syllable repetition in Martin et al. (2018) and response planning in Bögels et al. (2018)), it must also be pointed out that in these two experiments the N400 should capture prediction effects. According to a popular account of prediction (Pickering & Gambi, 2018), prediction involves preactivation of linguistic representations. Therefore, it is not clear whether speech planning would still affect conceptual processing when participants are not predicting during sentence comprehension. Another open question concerns whether conceptual processing during comprehension - indexed as N400 effects - is also affected when the comprehension task includes single words rather than full sentences. Since single words usually do not involve pre-activation of linguistic information, they provide a good study case to investigate how online comprehension is affected by linguistic dual-tasking in the absence of prediction processes. In the current experiment, we filled this gap by investigating any effects of linguistic dual-tasking on the auditory N1 and N400 when the comprehension task includes single words.

4.1.2 Effect of linguistic dual-tasking on repetition priming

While the studies in the previous section mainly focused on online comprehension, two previous studies looked at the effect of concurrent production and comprehension on how well comprehended words are encoded and stored for later use (Bartolozzi et al., 2021; Jongman & Meyer, 2017). In particular, they investigated whether speech planning hinders priming of comprehended words. It has been widely shown that the type of relationship between words (identical, associative, phonological, semantic etc.) can either facilitate or hinder processing (Abdel Rahman & Melinger, 2011; Belke & Stielow, 2013; Schriefers et al., 1990; Wheeldon & Monsell, 1992). Given that priming mechanisms play an important role in an influential model of conversation (interactive alignment model, Garrod & Pickering, 2009; Pickering & Garrod, 2004, 2013), it is important to determine whether they can also be used when comprehension is taxed.

Indeed, if speech planning disrupts online comprehension processes, the resulting linguistic representations might be degraded and less precise than those created during simple comprehension, leading to priming decrements.

In Jongman and Meyer (2017), participants heard a prime word and simultaneously saw a distractor picture for 250ms. Two seconds after distractor onset, participants saw a second picture, the target, that was either unrelated, identical, or associatively related to the prime. Participants were always asked to name the target picture. As for the distractor picture, participants were asked to either name it (Plan condition) or to listen to an auditory presentation of the picture's label (No Plan condition) in the interval between the distractor and target onsets. The planning condition was implemented between subjects in a first experiment and within subjects in a second experiment. In the first experiment, both repetition (identical) and associative priming were comparable across planning conditions; by contrast, in the second experiment, associative priming was greater in the no plan than in the plan condition, while repetition priming was similar across conditions.

A subsequent experiment by Bartolozzi et al. (2021) used a similar paradigm to the first experiment in Jongman and Meyer (2017), this time only focusing on repetition priming. Unlike Jongman and Meyer (2017), in this study primes were embedded in sentences, and prime-target pairs were separated by either 0, 10 or 50 trials. The authors used different lags to investigate whether linguistic dual-tasking impacted on the rate of decay of repetition priming across intervening trials. Participants always named the distractor pictures: half of the participants saw the distractor picture at the end of the prime sentence (Experiment 1, No Plan), while the other half saw the distractor picture at the onset of the prime word in the sentence (Experiment 2, Plan). Repetition priming did not differ across experiments, regardless of the distance between prime and target, suggesting again that repetition priming was resilient to divisions of attention.

The results from Jongman and Meyer (2017) and Bartolozzi et al. (2021) suggest that words comprehended with or without concurrent word planning elicit similar repetition priming effects. The lack of a behavioural effect is not surprising, given that implicit memory tasks such as priming can be quite resilient to divisions of attention (Mulligan, 1998; Spataro et al., 2011). However, repetition priming decrements during dual-tasking can occur if the secondary task is made more demanding, i.e., by increasing the frequency of the response to the distractor (Mulligan et al., 2007). In both Jongman and Meyer (2017) and Bartolozzi et al. (2021) the SOA between the onset of the distractor and of the target was about 2s. The lack of repetition decrements in the plan condition could be due to two different reasons: first of all, it is possible that there was no interference between speech planning and comprehension; alternatively, it is possible that some interference occurred but that the SOA between the distractor and the target made it possible for participants to recover and thoroughly process the primes. Since the paradigms used by Jongman and Meyer (2017) and Bartolozzi et al. (2021) did not make it possible to distinguish these two hypotheses, we tried to disentangle the two possibilities by developing a repetition priming paradigm where EEG activity was measured at the onset of the identical or unrelated spoken word. Any differences between No Plan and Plan conditions would suggest that dual-tasking affects initial encoding of a word, even if there is no behavioural effect. If the two conditions do not differ between each other, we conclude that word processing is resilient to any divisions of attention.

4.2 The current study

In this experiment, we used a modified version of Jongman and Meyer (2017) and Bartolozzi et al. (2021) to address two questions. While previous studies have shown that picture-naming can interfere with concurrent comprehension of tones and syllables, our main question was whether there is interference if the comprehension task includes meaningful words. As a side question, we also asked whether repetition priming effects are greater for words encoded during simple comprehension than for words encoded under linguistic dual-tasking.

The experiment was divided into two blocks (no plan and plan blocks), which were presented to participants consecutively. Both blocks included a series of alternating study and test phases. In the study phases, participants heard a prime word while naming a distractor picture (plan block, P) or passively listening to the name of a distractor picture (no plan block, NP). In the test phases, they named target pictures, which could be identical or unrelated to the primes in the preceding study phase. Unlike previous studies (Daliri & Max, 2016; Fargier & Laganaro, 2019), the comprehension task involved words rather than syllables, so as to be able to measure event-related responses to both early auditory processing and later processing of the meaning of the stimuli.

To answer our first question, i.e., whether linguistic dual-tasking impacts on online comprehension processes, we analysed data from the study phases in both the NP and P blocks. We hypothesised that the amplitude of the N1 (indexing early auditory processing) and the amplitude of the N400 (indexing later processing) might be affected by linguistic dual-tasking. More specifically, we predicted that the amplitude of the N1 and N400 would be greater in the NP block than in the P block, due to the fact that participants can fully allocate their attention to the comprehension task, without interference from speech planning processes. Related to this, we tried to replicate previous findings of interference of comprehension on picture-naming: in particular, we tested whether the amplitude of the N1 and N400 could be predicted by the distractor naming latencies, which would again suggest a trade-off between comprehension and production.

In order to answer our second research question, that is to say whether there is an effect of linguistic dual-tasking on repetition priming, we also measured EEG activity when participants planned to produce the primed word. Our aim was to compare behavioural and EEG responses in the test phases of both the NP and P blocks. In order to avoid articulation-related artifacts while recording EEG in the test phase, the distance between the onset of the target picture and the moment participants could speak was quite long (1s). One empirical question was therefore whether it is possible to still detect priming effects despite such a long delay before articulation. Following Jongman and Meyer (2017) and Bartolozzi et al. (2021), we predicted that any repetition priming effects, measured as shorter target naming latencies for identical than unrelated words, would not differ across the test phases of the NP and P blocks. For this reason, we did not expect to find any modulation of planning condition on the N400 effect for identical vs unrelated targets. Identical targets typically yield a reduced N400 with respect to unrelated targets (Kutas & Federmeier, 2011). If linguistic dual-tasking impacts on repetition priming - even if no behavioural modulation occurs - the N400 reduction following identical primes should be greater during single-tasking than during dual-tasking. By contrast, if the reduction does not differ between the conditions, we would conclude that linguistic dual-tasking has no effect on repetition priming.

While Jongman and Meyer (2017) and Bartolozzi et al. (2021) did not find any modulations of planning condition on repetition priming effects, Fargier and Laganaro (2019) and Bögels et al. (2018) showed that participants might prioritise comprehension or production on a trial-by-trial basis, yielding trade-offs between production and comprehension. For this reason, we aimed to measure whether the repetition priming effect in both the NP and P blocks correlated with the N400 elicited by the prime word. The hypothesis was that, if there is a trade-off between production and comprehension, there should be a negative correlation between the target naming latencies of identical targets in the test phase and the N400 effect of the corresponding primes in the study phase.

4.2.1 Participants

Fourty-one participants (23 female, mean age = 24.95, range = 19 - 35) took part in the study. They were all native speakers of Dutch with normal-to-corrected vision and no speech disorders. One of them was left-handed and one did not provide information about handedness. All participants were recruited using the Max Planck Participant Database, provided verbal and written consent prior to participating in the experiment, and received € 18 as compensation. This study obtained ethical approval by the Faculty of Social Sciences of the Radboud University.

4.2.2 Methods

Materials

The experiment was divided into two blocks: a no plan (NP) block and a plan (P) block. Each block included 74 trials. Therefore, we created four lists, each containing 74 trials, which could be presented in either the NP or P block. The condition in which the prime-target pairs occurred (identity or unrelated) was counterbalanced across lists. In each list, half of trials were identical trials, the other half were unrelated trials. Each participant was presented with two lists: one in the NP block and the other in the P block. The lists were counterbalanced across participants.

Pictures. All pictures were selected from the Multipic database (Duñabeitia et al., 2018). 148 target pictures were selected to be used as identical targets in the test phases. Each target picture was matched with an unrelated distractor picture and two spoken primes, an identical prime and an unrelated prime (see Appendix A for a comprehensive overview of the materials). Target pictures had a mean name agreement of 93.24% (range: 52.63% - 100%) and a mean frequency (fpm, frequency per million in the SUBTLEX database, Keuleers et al., 2010) of 8.54 (range: 0.02 – 44.07). We also selected 148 pictures to be used as distractors in the study phase. These pictures had mean name agreement of 84.49% (range: 38.60% – 100%) and mean fpm of 121.03 (0.07 – 4412.02).

Spoken primes and distractors. A native speaker of Dutch recorded the sound files in a shielded booth. We recorded 370 words. After being recorded, words were segmented in separate files and normalised using Praat (version 6.1.09).

148 words were used as identical primes, and 74 words (37 in each block) were used as unrelated primes (mean fpm: 8.86, range: 0.07- 28.52). Each unrelated prime was coupled with two targets (one presented in the identical condition and one in the unrelated condition in each list, so that the unrelated prime occurred once throughout the experiment). 148 other words matched the distractor pictures and were played in the study phase of the NP block (as described above, in this block participants did not name the distractor pictures but rather listened to their names). All materials, datasets and scripts used for the analyses will be freely available on the Open Science Framework upon publication.

Procedure

The experiment was divided into two blocks, a NP block and a P block. The P block always followed the NP block to minimize the likelihood that participants would covertly name the distractor pictures in the NP block. Each block was divided into five study phases and five test phases, which alternated throughout the block. Each study and test phase included fifteen trials, with the exception of the last study and test phases of each block, which included 14 trials. Primes and distractors were presented in the study phase, while targets were presented in the test phase. The unrelated/identical primes in the study phase always referred to the targets presented in the immediately following test phase.

Each trial in the study phase began with the presentation of a fixation cross, with a random duration between 450 and 550ms. The fixation cross was followed by a distractor picture ("gitaar", *guitar* in Figure 4.1), which stayed on the screen for 350ms, and by a prime word ("zaklamp", *torch*), which was played 150ms after the onset of the distractor picture. The distractor picture was then followed by a fixation cross for 1300s, and by a second fixation cross (NP block) or exclamation mark (P block) for 1.5s. In the NP block, the name of the distractor picture ("gitaar") was played at the onset of the second fixation cross. In the P block, the exclamation mark signaled that participants could name the distractor picture. Each trial ended with the presentation of a mask (####) for 2s. Immediately after each study phase, the corresponding test phase began. A fixation cross was shown for a variable duration, between 450 and 550ms. It was followed by a target picture, which stayed on the screen for 350ms. The name of the target picture could be identical to one of the primes presented in the immediately preceding study phase, or it could be an unrelated name. The target picture was followed by a fixation cross for 650ms and by an exclamation mark for 1.5s. The exclamation mark signaled to participants that they should name the target picture aloud. Each trial in the test phase ended with the presentation of a mask (####) for 2s.

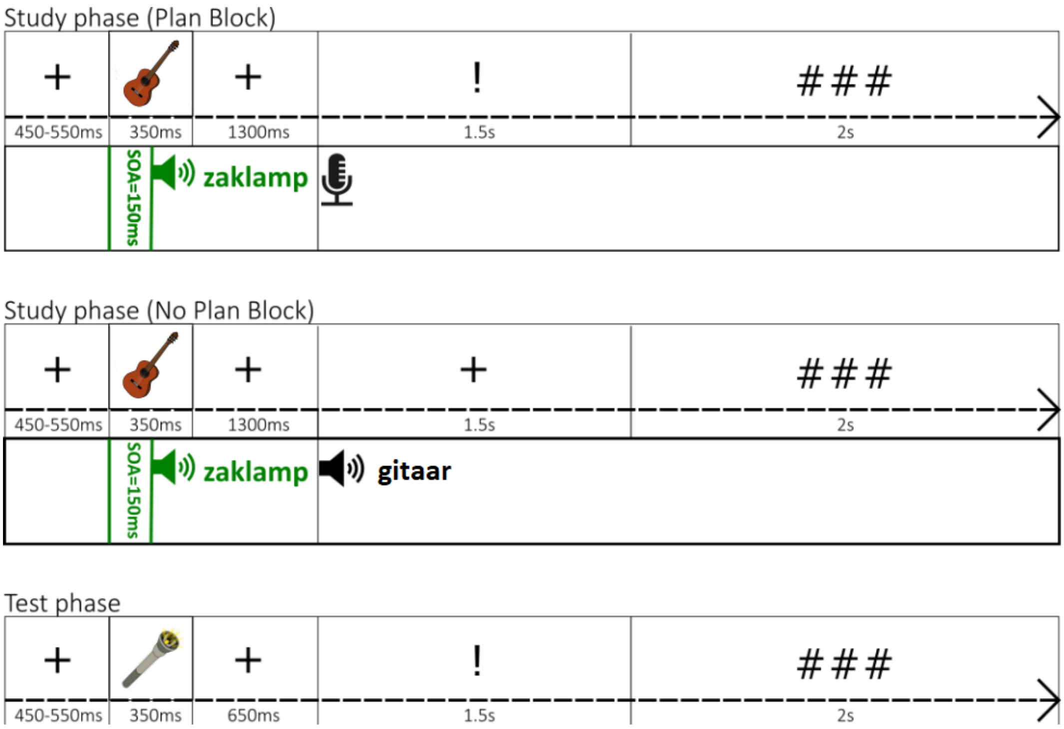

Figure 4.1: Example of a trial with identical prime and target in the study phase (Plan and No Plan) and in the test phase. The numbers in black indicate the duration of each item on the screen.

Data acquisition

We recorded the EEG using an EasyCap montage and silver-silver chloride electrodes. The cap included 61 electrodes, placed following the international 10-10 system. Two electrodes were placed on the left and right mastoids, and one below the eye to record eye blinks. The electrodes were referenced online to the electrode on the left mastoid. The EEG data were recorded using a BrainAmp DC amplifier (Brain Products GmbH, Gilching, Germany) at a sample rate of 500 Hz, and filtered online with a high cutoff filter of 249Hz and a low cutoff filter of 0.016Hz. Electrode impedance was kept below 25 kΩ. The experiment was administered using Presentation, version 18.3 (https://www.neurobs.com/).

Data preprocessing

Study phase. EEG preprocessing and analyses were carried out in MATLAB (R2020b; Mathworks, Inc.) using the Fieldtrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011). The EEG recordings were initially low-pass filtered at 100 Hz and high-pass filtered at 0.1 Hz using a windowed sinc FIR filter. In order to account for the interference from power lines, the data were also band-stop filtered at 50 Hz, 100 Hz, and 150 Hz. The data were first segmented into epochs ranging from 850ms before distractor picture onset to 1650ms after distractor picture onset. The data were re-referenced to the average of the left and right mastoid and baseline corrected using a window between -200ms and 0ms before the onset of the distractor picture. The data were inspected for artifacts in three different steps: first, we removed extreme outliers following visual inspection; we then performed ICA (*runica* implementation in Fieldtrip) on a filtered version of the data (high-pass filter at 1Hz), and then applied the unmixing matrix to the non-filtered version of the data. We removed components related to eye blinks, eye movements and heart activity (mean number of removed components per participant = 1.6, range = 1-4), and then recombined the remaining components. To check whether the data still contained any eye movement-related activity, we detected trials based on voltage threshold (cutoff: 100μV) and step (cutoff: 35μV), and removed them manually. After preprocessing the data, we inspected the waveforms to check for any remaining artifacts and calculated the number of trials that were excluded after artifact rejection. We excluded datasets with less than 60% of trials remaining.

Prior to the statistical analyses, we re-epoched the data according to the onset of the prime audio (occurring 150ms after distractor onset), and created epochs ranging from -600ms to 1500ms with respect to prime onset. We then low-pass filtered the data at 30Hz, subset correct trials, and split them according to block (NP vs P). Since no response was required in the NP block, no correct responses could be subset. As a result, the number of trials in the P block was smaller than in the NP block. Since this could lead to different signal-to-noise ratios in the two blocks, potentially making any comparisons problematic, we equated the number of trials in the two blocks by taking a random subsample of trials from the NP block. We then applied a baseline correction using a 200ms window before distractor picture onset (picture onset = 150ms before prime onset), averaged trials within each condition and block, and computed grand average waveforms.

Test phase. The EEG preprocessing steps applied to the data in the test phase were the same as those used in the study phase, except that in the test phase data were segmented into epochs ranging from -1350ms to 650ms after the target picture offset and that the 200ms baseline window was taken with respect to the onset of the target picture. The number of ICA components removed in the test phase datasets ranged from 1 to 8 (average: 1.9).

Before running the statistical analyses, we re-epoched the data (from -1s to 1s with respect to target picture onset) and applied a low-pass filter at 30Hz. We then subset correct trials and split them according to block (P versus NP), and target relatedness (identical versus unrelated). We then applied a 200ms baseline correction with respect to target onset, averaged trials within each block and condition and computed grand average waveforms.

Statistical analyses

Behavioural data. The behavioural analyses reported here included 40 participants (one participant was excluded due to technical problems during the experiment). The behavioural analyses including only the participant datasets that were used in the EEG analyses (n=30, see section below) are reported in Appendix B.

Accuracy and picture-naming latencies were modelled using generalised linear mixed effect models and linear mixed effect models, respectively (package lme4, version 1.1-23, Bates et al., 2014) in R (version 4.0.2, Team, 2020). Prior to the analysis, latencies were centred and log-transformed. All predictors were contrast coded using sum-to-zero contrasts. To determine the random-effects structure, we first fitted the model using the maximal random-effects structure. We then simplified the random-effects structure according to Bates et al. (2015). To determine whether adding a predictor improved model fit, we ran Likelihood Ratio Tests of the models with the predictor and without the predictor of interest. For glmer models, we also calculated 95% confidence intervals using confint (method = "profile").

EEG data. A cluster-based permutation approach (Maris & Oostenveld, 2007) was used to detect any differences between the conditions of interest. Analyses were carried out on a window between 0ms and 1000ms after prime onset for the study phase data, and on a window between 0ms and 800ms after target picture onset for the test phase data. We initially used a dependent samples t-test (uncorrected alpha level = 0.05) to determine data points that showed an effect for the contrast of interest. Data points with alpha <.05 were then used to calculate clusters in the data, while all other data points were excluded. We then calculated the sum of the t-values of the data points in a cluster (cluster-based statistics). We then permuted participants' averages between the conditions of interest 5000 times and calculated again a cluster-based statistic. We created a permutation distribution by selecting the largest statistic during each iteration. If the original cluster-based statistics exceeded alpha level = 0.05 (two-sided), we rejected the null hypothesis.

4.2.3 Behavioural results

	Distractor accuracy rate and mean naming latency		
Relatedness	Accuracy rate	Mean Naming latency (ms)	SD (ms)
Identical	0.72	419	175
Unrelated	0.72	414	168

Table 4.1: The table includes distractor accuracy rates and naming latencies (mean and standard deviation) by Relatedness in the study phase.

Study phase. Accuracy rates and mean naming latencies in response to distractor pictures are summarised in Table 4.1. We investigated whether distractor accuracy rates and naming latencies in the study phase were predicted by relatedness. This analysis was carried out to ensure that any differences between unrelated and identical trials in the test phase were not due to factors unrelated to priming. This means that relatedness should not predict accuracy rates and naming latencies in the study phase. Both models included relatedness as predictor, and by-participant and by-distractor intercepts. The analyses confirmed this hypothesis: indeed, adding relatedness to the model as a predictor did not improve model fit with respect to the null model (accuracy: $\chi^2(1) = 0.002$, $p = .969$; naming latencies: $\chi^2(1) = 0.79$, $p = .373$).

Test Phase. Accuracy rates and mean naming latencies in the test phase are reported in Table 4.2.

		Target accuracy rate and mean target naming latency		
Relatedness	Block	Accuracy rate	Mean Naming latency (ms)	SD (ms)
Identical	*Plan*	0.89	446	174
Identical	*No Plan*	0.92	488	201
Unrelated	*Plan*	0.86	462	190
Unrelated	*No Plan*	0.88	494	203

Table 4.2: The table includes target accuracy rates and naming latencies (mean and standard deviation) by relatedness and block in the test phase.

In the first analysis, we used mixed-effects logistic regression to determine whether accuracy could be predicted by relatedness (identical vs unrelated), block (plan vs no plan), and their interaction. While relatedness ($\chi^2(1) = 25.85$,

$p = <.001$) and block ($\chi^2(1) = 8.73$, $p = .003$) predicted accuracy rates, their interaction did not ($\chi^2(1) = 2.74$, $p = .098$). The final model, together with confidence intervals, is summarised in Table 4.3.

	Fixed effects						Random effects		
Fixed	Estimate	SE	z	p.value	lower_CI	upper_CI	Random	Variance	sd
Intercept	2.69	0.13	20.00	<0.001	2.43	2.98	Participant (Intercept)	1.53	1.24
Target relatedness	-0.23	0.05	-5.20	<0.001	-0.33	-0.14	Target picture (Intercept)	0.15	0.38
Block	-0.14	0.04	-3.04	0.002	-0.23	-0.05			

Table 4.3: Mixed effects model with target accuracy as the dependent variable, and target relatedness and block as the independent variable.

In the second analysis, we determined whether relatedness, block and their interaction affected naming latencies. The only significant predictor was block ($\chi^2(1) = 8.26$, $p = = .004$), while relatedness ($\chi^2(1) = 2.6$, $p = .107$) and interaction ($\chi^2(1) = 2.3$, $p = .130$) had no effect. The final model is summarised in Table 4.4.

	Fixed effects			Random effects		
Fixed	Estimate	SE	t	Random	Variance	sd
Intercept	-0.002	0.04	-0.06	Target picture (Intercept)	0.013	0.12
Block	-0.032	0.01	-3.03	Target Picture (Block)	0.002	0.04
				Participant (Intercept)	0.054	0.23
				Participant (Block)	0.003	0.06

Table 4.4: Mixed effects model with log-transformed naming latencies as the dependent variable, and block as the independent variable

The small priming effect in the accuracy analysis indicates that participants occasionally benefitted from previously having heard the picture name. The small effect size might be due to the fact that target pictures had medium to high name agreement. Therefore, participants tended to use the same name for most of the pictures regardless of any priming effects and repetition benefits only occurred for a few items.

As for the naming latencies, the lack of priming is not unexpected: indeed, the test phase included a delayed naming task, in which participants were asked to name the target picture 1s after its onset. Given that picture naming typically takes about 600ms (Indefrey & Levelt, 2004), it is possible that, in the current experiment, participants ended production processes well before they could start speaking, and they withheld their response until they were allowed to speak, therefore washing away any naming latency differences between identical and unrelated targets.

4.2.4 ERP results

While we tested 41 participants, the analysis only included data from 30 participants. One participant was excluded due to technical problems during the experiment, one participant was excluded because too many trials were rejected in the preprocessing stage, and nine participants were excluded due to poor quality of the data (muscle activity and/or excessive drift).

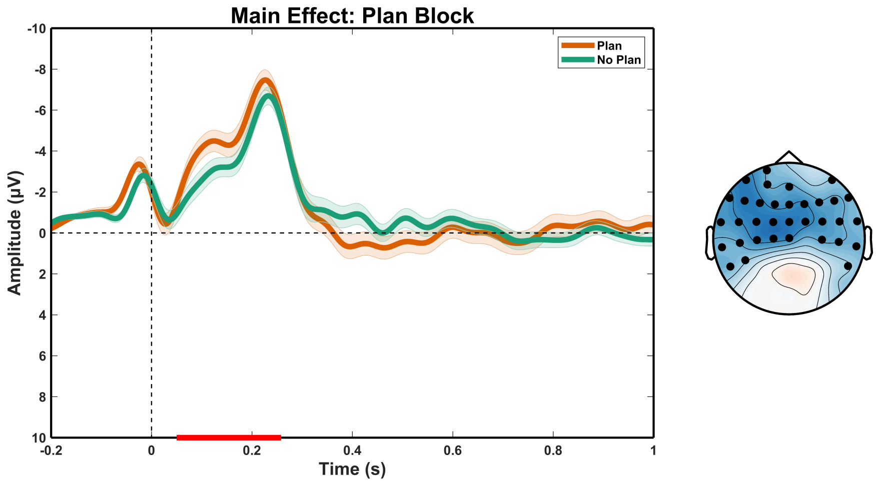

Figure 4.2: First significant cluster in the study phase. The figure depicts the first cluster in the study phase (difference between P and NP blocks). *Left*: waveforms describing the time course in the NP and P blocks (green = NP, orange = P), averaged over the electrodes contributing to the cluster. The shaded area around the waveforms indicates the standard error of the mean. The significant cluster is indicated by the red line on the x axis. *Right*: scalp topographies describing the difference between the P and NP blocks in the time window corresponding to the significant cluster. The filled circles show electrodes from the cluster contributing at least 50% of time interval.

Study phase. We compared trials in the P and NP blocks in which participants correctly named the distractor picture. The analysis yielded two statistically significant clusters (see Figures 4.2 and 4.3): a first cluster with a fronto-central distribution ($p = .01$), and a second cluster with a centro-parietal distribution ($p = < .001$). Based on the fronto-central topography, and negative peak latency around 100ms, we identify the first cluster as an N1 ERP component (Luck & Kappenman, 2011). While we initially predicted that the N1 would have a greater amplitude in the NP than in the P block, the data showed the opposite pattern, i.e., greater N1 amplitude in the P block than in the NP block. In a subsequent analysis, we also tested whether the amplitude of the N1 in the P block could be modulated by the distractor naming latencies.

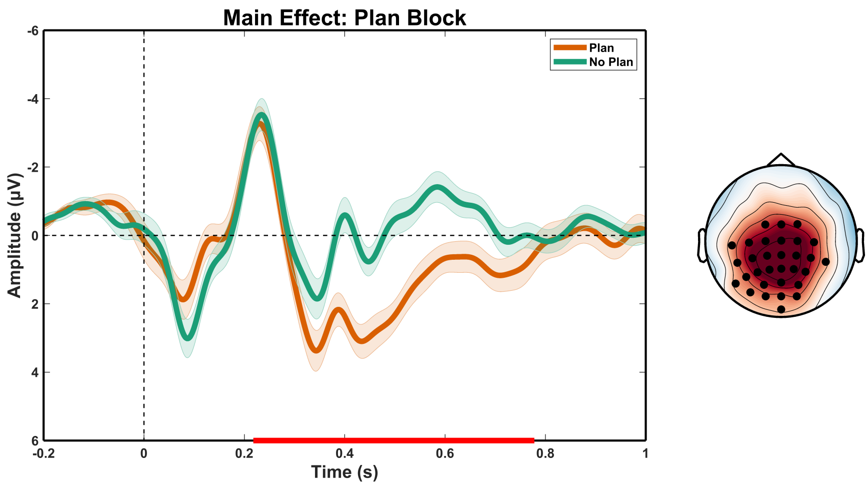

Figure 4.3: Second significant cluster in the study phase. The figure depicts the second cluster in the study phase (difference between P and NP blocks). *Left*: waveforms describing the time course in the NP and P blocks (green = NP, orange = P), averaged over the electrodes contributing to the cluster. The shaded area around the waveforms indicates the standard error of the mean. The significant cluster is indicated by the red line on the x axis. *Right*: scalp topographies describing the difference between the P and NP blocks in the time window corresponding to the significant cluster. The filled circles show electrodes from the cluster contributing at least 50% of time interval.

Given previous studies that found a trade-off between comprehension and production, we argued that faster responses to the distractor picture might be associated with a decreased N1 and/or N400 after the onset of the prime. This would suggest that, on trials where participants shifted their attention to the production task, they were faster to name the distractor picture. However, this might come at the expense of the comprehension task, as indexed by reduced N1 and N400 components in response to the auditory prime. We therefore ran two linear mixed-effects models. The first one included by-participant and by-trial N1 amplitude as the dependent variable, and the log-transformed distractor naming latencies as the predictor. The random-effects structure included by-participant and a by-target intercepts. Contrary to our predictions, this model did not perform better than the null model ($\chi^2(1) = 1.37$, $p = .241$). We then ran an identical model, where the dependent variable was the by-participant and by-trial N400 amplitude. As in the previous model, adding the distractor naming latencies as predictor did not improve model fit ($\chi^2(1) = 0.23$, $p = .632$).

As for the pattern shown in the second cluster, we advance two hypotheses. A first possibility is to interpret the difference between conditions as a reduced

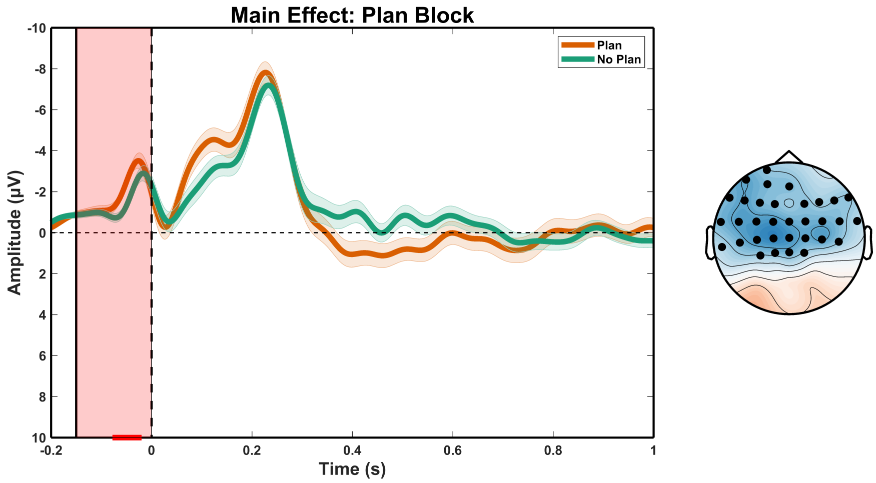

Figure 4.4: Significant cluster in the time window before the onset of the prime word (-150ms - 0ms). *Left*: waveforms describing the time course in the NP and P blocks (green = NP, orange = P), averaged over the electrodes contributing to the cluster. The analysed time window is highlighted in pink. The shaded area around the waveforms indicates the standard error of the mean. The significant cluster is indicated by the red line on the x axis. *Right*: scalp topographies describing the difference between the P and NP blocks in the time window corresponding to the significant cluster. The filled circles show electrodes from the cluster contributing at least 50% of time interval.

N400 in the P block with respect to the NP block, in agreement with our prediction. While the topography of the effect observed in the data is compatible with an N400, the waveform shows an atypical morphology. Indeed, the N400 is usually described as a negative-going waveform peaking between 300ms and 500ms after stimulus onset. However, the waveforms in this experiment show a much more sustained difference, which seems to last at least until 700ms post stimulus onset. This, together with the fact that the waveforms show positive-going - rather than negative-going - deflections, suggests that they could represent a sustained positivity, which is enhanced in the P block relative to the NP block. We will return to this point in the General Discussion.

Since Figure 4.2 also shows a possible difference before the onset of the prime, we decided to run further analyses. This difference prior to the prime could be induced by the distractor picture, which appeared on the screen 150ms before the prime word onset. We therefore decided to run a cluster-based permutation test in the time window from -150ms to 0ms before prime onset (i.e., the time period between the onset of the distractor picture and the onset of the prime). It is important to keep in mind that this time period does not correspond to the baseline window, which is between -350ms and -150ms relative to the onset of the prime (and therefore precedes the onset of the distractor picture). The results showed a negative cluster ($p = .004$) with a fronto-central topography (see Figure 4.4). We hypothesised that this difference might reflect a delayed visual N1 onset in the NP compared to the P block. This difference might reflect a practice effect induced by the fixed block order (the P block always follows the NP block) or task demands. Since participants named the distractor in the P block but not in the NP block, processing of the distractor picture may be more thorough in the former than in the latter.

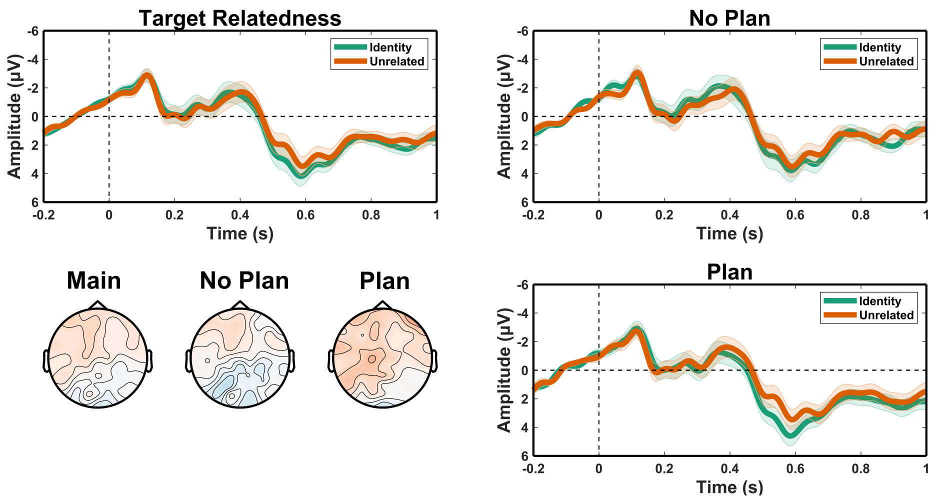

Figure 4.5: Identical vs unrelated trials in the test phase. The figure depicts the difference between identical unrelated correct trials in the test phase (no statistically significant effects were found). *Top Left*: waveforms describing the time course in the unrelated and identical trials (green = identity, orange = unrelated), averaged over electrodes C1, Cz, C2, CP1, CPz, CP2, P1, Pz, P2, CP3, CP4. The shaded area around the waveforms indicates the standard error of the mean. *Top right and bottom right*: waveforms split by block. *Bottom left*: scalp topographies describing the difference between the identical and unrelated trials in the typical N400 time window (between 300ms and 500ms post stimulus onset). The first topography describes the difference between identical and unrelated trials across blocks, while the second and third plot describe the difference between identical and unrelated trials split by block.

Test Phase. We carried out three analyses on the data of the test phase: first, we compared unrelated and identical trials (main effect of relatedness), then trials in the NP and P blocks (main effect of block), and then the difference between unrelated and identical trials in the NP and P blocks (a way of testing for an interaction in the cluster-based framework). While we did not find any

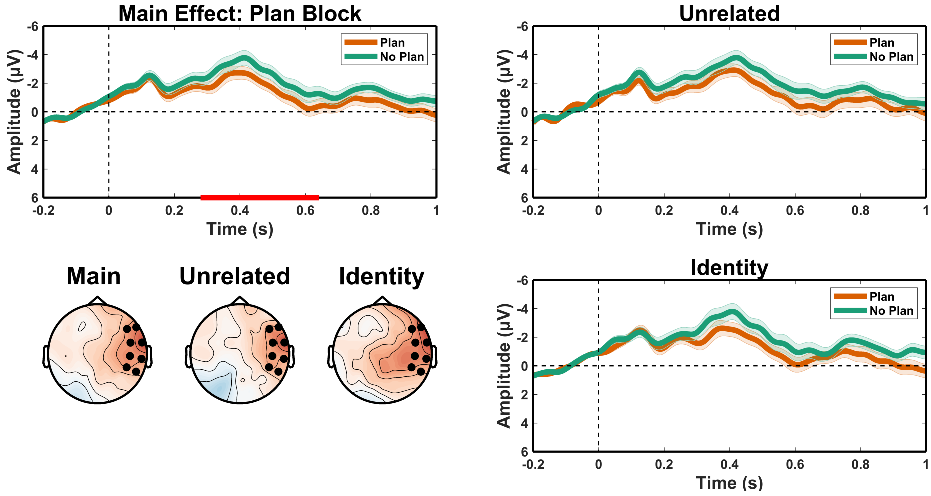

Figure 4.6: Significant cluster in the test phase. The figure depicts the significant cluster in the test phase (difference between P and NP blocks). *Top Left*: waveforms describing the time course in the NP and P blocks (green = NP, orange = P), averaged over the electrodes contributing to the cluster. The shaded area around the waveforms indicates the standard error of the mean. The significant cluster is indicated by the red line on the x axis. *Top right and bottom right*: waveforms split by relatedness. *Bottom left*: scalp topographies describing the difference between the P and NP blocks in the time window corresponding to the significant cluster (the first plot describes the main effect of block, while the second and third plot describe the effect of block split by relatedness. The filled circles show electrodes from the cluster contributing at least % of time interval.

effects of relatedness or of the interaction between block and relatedness ($p >$ .05, see Figure 4.5), the comparison between P and NP blocks yielded a positive cluster ($p = .02$, see Figure 4.6). Given that the latter comparison yielded an atypical topography and waveform, we inspected by-participant topographies to determine whether the effect was driven by participants with extreme values in the time window and at the electrodes detected by the cluster. We then identified three participants that could drive the effect and re-ran the preprocessing of these datasets to check for any remaining artifacts. Since the results did not change after the second round of artifact rejection, we re-ran the cluster based permutation test without these participants. When we excluded these participants, the cluster was no longer significant (See Appendix B).

4.3 Discussion

Using both behavioural and electrophysiological measures, we asked whether linguistic dual-tasking has an impact on comprehension processes. Our primary focus was whether concurrent speech planning interferes with online comprehension of single words. We tested this research question in the study phase of the current experiment, during which participants heard words while planning the name of a picture (P block) or without any additional tasks (NP block). Given previous studies on this topic, we advanced three hypotheses concerning the electrophysiological measures. First, we hypothesised that the N1 in response to an auditory word would be reduced or delayed in the P block than the NP block, due to the fact that, in the P block, participants had to split their attention between comprehension and production, possibly leading to interference. Similarly, we expected the N400 in response to the prime word to be smaller in amplitude during dual-tasking than in single-tasking. Third, we anticipated a trade-off between comprehension and production, whereby faster naming latencies in the production task would be associated with a reduced N1 and N400 in the comprehension task.

As a second research question, we also investigated whether re-activation and retrieval of a word, indexed as priming effects for identical vs unrelated targets, could be influenced by whether the word had initially been encoded during dual-tasking or single-tasking. We evaluated this hypothesis in the test phase of the current experiment, during which we presented participants with pictures whose labels had been or had not been used as words in the comprehension task of the study phase. Given previous studies (Bartolozzi et al., 2021; Jongman & Meyer, 2017), we did not expect to find any behavioural differences between the size of priming of words initially encoded in silence compared to during speech planning. While we did not expect any behavioural differences, we hypothesised that relevant differences would emerge in the EEG signal. In particular, we expected the N400 effect for identical vs unrelated words to be greater in the NP block than in the P block. However, in order to ensure that the EEG data recorded in the test phase were not contaminated by speech-related artefacts, we had to use an unusually long SOA between the onset of the target picture and the moment participants could speak. Therefore, it was an empirical question whether any priming effects (indexed as shorter naming latencies for identical than unrelated trials) would occur at all in the analysis of picture-naming latencies.

4.3.1 Effect of linguistic dual-tasking on online comprehension processes

The results of the EEG analysis in the study phase did not support our hypothesis. First of all, the cluster-based permutation yielded a negative cluster over fronto-central electrodes. We identified this peak as an N1 elicited by the prime word. While we predicted that the N1 would be greater in the NP than P block, we found the opposite pattern, i.e., a reduced N1 in the NP block with respect to the P block. The analysis also yielded a second cluster, showing a positive difference over centroparietal electrodes in the comparison between the P and NP block. We advanced two hypotheses as to how to interpret this effect: either as a reduced N400 in the P block than in the NP block, or as an enhanced sustained positivity in the P block than the NP block.

In dichotic listening studies, the amplitude of the N1 is reduced when the stimulus is presented in the unattended rather than the attended channel (e.g., Hillyard, Hink, Schwent, & Picton, 1973; Parasuraman, 1978; Schwent, Hillyard, & Galambos, 1976), and similar results have been obtained in cross-modal attention studies (e.g., Luo & Wei, 1999; Parasuraman, 1985). As a result, we predicted that, in our experiment, the N1 would be reduced in the P block with respect to the NP block. The hypothesis was that, in the P block, the attention would be split between the distractor-naming task and the comprehension task. This should lead to a decrement in the N1 amplitude with respect to the NP block, where attention can be fully allocated to the comprehension task. However, we found the opposite pattern, i.e., a reduced N1 in the NP than P block.

In our experiment, the decreased N1 in the NP relative to the P block could depend on specific order effects. It is possible that participants started paying more attention to the auditory words when they noticed that some of those corresponded to the picture names in the subsequent test phase. Focused attention is known to increase the amplitude of the auditory N1 (e.g., Hink, Van Voorhis, Hillyard, & Smith, 1977). Since the P block always followed the NP block, this increase of attention towards the prime word might be more evident in the former than the latter, therefore yielding the N1 pattern above.

A possible objection to this explanation is that previous studies have found a decreased/delayed N1 during dual-tasking compared to single-tasking even when the comprehension task only involved passive hearing, i.e., the task did not require a response to the auditory input (Daliri & Max, 2016; Fargier & Laganaro, 2019). This means that some reduction of the N1 during dual-tasking rather than single-tasking should still be evident in the current experiment, regardless of any specific-order effects. However, previous studies found effects of linguistic dual-tasking on the auditory N1 only when the SOA between the production and the comprehension task was much longer than in our experiment (>= 400ms vs 150ms). As described in the introduction, Fargier and Laganaro (2019) did not find any delayed onset of the auditory N1 at SOA=150ms, which is compatible with the pattern of results obtained in our experiment. In addition to this, in both Daliri and Max (2016) and Fargier and Laganaro (2019) the auditory stimulus only occurred in a subset of the trials (40% in Daliri and Max (2016) and 75% in Fargier and Laganaro (2019)), meaning that in some of the trials participants saw pictures but did not receive any auditory input. By contrast, in this experiment the heard word occurred in all trials and all blocks, possibly washing away any N1 effects due to the occurrence/absence of the auditory input.

In addition to this first negative cluster, the analysis yielded a second positive cluster. We provide two possible explanations for this pattern of results. A first possibility is that this cluster describes a reduced N400 in the P block with respect to the NP block, which would be consistent with our initial prediction. The N400 has been often interpreted as a measure of semantic processing (e.g., Kutas & Federmeier, 2000; Lau, Phillips, & Poeppel, 2008). Following this interpretation, speech planning would not interfere with early comprehension processes, as indexed by the fact that the N1 was reduced in the NP rather than in the P block (contrary to our predictions), but would affect later semantic processing of meaningful stimuli.

However, as discussed in the results section, the morphology of the waveform is not that of the typical N400, but is more consistent with a sustained positivity. Such positivity has previously been found in studies investigating comprehension during linguistic dual-tasking (Bögels et al., 2018; Gerakaki, 2020), but also speech planning during comprehension (Bögels et al., 2015), and immediate and delayed picture-naming (Eulitz, Hauk, & Cohen, 2000; Jongman, Piai, & Meyer, 2020). Since this effect has been found in studies that did not involve any comprehension tasks, it likely reflects processes related to the production task. In our paradigm, this positivity could reflect four different mechanisms. First, it could reflect working memory demands because we employed a delayed-naming task, in which participants had to keep the picture name in working memory until they could speak. Second, it could index speech planning processes, as suggested by Bögels et al. (2015). Third, it could reflect increased attention to the target and, lastly, general decision processes or response preparation.

While these are all likely explanations of our pattern results, it might be possible to narrow down the number of hypotheses by considering previous studies that found the same pattern of results. An explanation in terms of working memory seems unlikely, since this positivity has also been found in studies using an immediate picture-naming task, in which there was no delay between the onset of the picture and the moment participants were allowed to speak (e.g., Eulitz et al., 2000). An interpretation in terms of speech planning processes is also disfavoured, since the positivity has also been found when participants are asked to perform a button-press task rather than give a spoken response (Jongman et al., 2020).

A more likely hypothesis is that the positivity reflects enhanced attention to the to-be-named picture. More specifically, Jongman et al. (2020) suggested that it might reflect "*attention to the final stimulus, the cue necessary to launch the response, but not the actual response planning itself*" (p.929). This account explains why this positivity has been found in immediate and delayed picture-naming tasks, and in tasks requiring a motor execution rather than a spoken response. In our experiment, participants might have monitored the occurrence of the distractor picture more carefully in the P block than the NP block, given that a response was only required in the former but not in the latter. A last possibility is that the positivity reflects either general decision processes (i.e., to respond or not) or response preparation. At this stage, it is not possible to distinguish between these hypotheses (attention, decision processes or response preparation).

It is important to point out that interpreting the difference between the P and the NP blocks as a positivity does not rule out the possibility that production may affect comprehension. Indeed, it is possible that the positivity is so prominent that it overrides other effects (e.g., N400 differences between blocks). This would point to serious limitations in the use of ERPs to investigate linguistic dual-tasking, at least in conjunction with paradigms similar to the one used in this experiment.

Unfortunately, the current paradigm does not enable us to decide whether the difference between the NP and P blocks in the second cluster reflects an N400, a positivity, or a combination of the two. If the difference reflects an N400, we would conclude that production impacts on comprehension. If the difference solely reflects a sustained positivity, this would suggest that concurrent production does not impact on comprehension. By contrast, an explanation in terms of a combined N400 and sustained positivity would suggest that production may affect comprehension but that the paradigm does not enable us to test this hypothesis because any N400 differences might be overridden by the sustained positivity.

To distinguish between these two hypotheses one could replicate the experiment using tones or syllables - instead of words - in the comprehension task. If the analysis yields a similar pattern of results to that obtained in this experiment, one would conclude that the positive cluster solely reflects a positivity. Indeed, tones should not elicit any N400 and such modulations should therefore not be evident in the data. Another possibility would be to add a picture-naming only task and a comprehension only task, so as to be able to compare how the waveforms differ in single-tasking and dual-tasking for each task (comprehension and production).

A last point of discussion concerns the absence of a trade-off between production and comprehension in the P block. We predicted that faster distractor naming latencies would be associated with a smaller N1 and/or N400, which would suggest that the degree of interference depends on the extent to which attention is directed toward production compared to comprehension. Correlations between naming latencies and the EEG signal in comprehension tasks have been found in previous studies (Bögels et al., 2018; Daliri & Max, 2016; Fargier & Laganaro, 2019). However, the data did not confirm this hypothesis, since neither the amplitude of the N1 nor that of the N400 were predicted by the distractor naming latencies.

It is possible that the lag between the onset of the distractor picture and the moment participants could speak (1650ms) was too long to capture any correlations between picture-naming latencies and the amplitude of ERP components. Due to the delayed response, the naming latencies may not reflect the stages of speech planning in the way they would in typical picture-naming studies. As a result, any interference between comprehension and speech planning is not reflected in the distractor naming latencies because participants need to wait before responding and have time to recover from any interference. Therefore, any differences in the ERP amplitudes might not correlate with the distractor-naming latencies because the latter do not reflect the amount of time necessary to plan the response.

Alternatively, the lack of any correlations in the current study might be due to the SOA between the onset of the distractor picture, indicating the beginning of the production task, and that of the prime word, i.e., signalling the comprehension task. Indeed, Fargier and Laganaro (2019) (and Daliri and Max (2016)) found interference using an SOA > 400ms, but found no interference at the SOA used in this experiment (150ms). As Fargier and Laganaro (2019) suggested, the amount of interference between comprehension and production may depend on the processes that are overlapping (e.g., lexical, post-lexical processes) and on the amount of allocated resources. Future studies investigating the effect of dual-tasking on comprehension of single words should use a wider range of SOAs, so as to determine exactly when interference occurs and which processes are affected.

4.3.2 Effect of linguistic dual-tasking on repetition priming

Our initial prediction was that identical targets would yield an advantage over unrelated targets, as indexed by higher accuracy rates and/or faster naming latencies. The results showed that repetition priming effects in the test phase were evident in the analysis of accuracy rates but not in that of naming latencies. Although significant, the effect of relatedness on target accuracy rates was quite small (the accuracy benefit for identical vs unrelated targets was 3% in the P block and 4% in the NP block). This is probably due to the fact that most target pictures in the experiment had mid to high name agreement, meaning that participants named these pictures with their modal name, regardless of priming. As a consequence, item repetition may have only been beneficial for a few items, leading to a small effect.

As for the analysis of picture naming latencies, we did not find any effect of relatedness. As specified in the Introduction, it was an empirical question whether the paradigm used in the test phase could capture repetition benefits on the naming latencies. In fact, the experiment involved a delayed-naming task, where the distance between the onset of the target picture and the time participants could speak was quite long (1s). Typically, picture naming takes about 600ms (Indefrey & Levelt, 2004). It is possible that, in this experiment, participants finished planning the name of the target before they were allowed to speak, therefore washing away any priming effects. While it was not possible to measure any repetition benefits on naming latencies, the priming effect obtained in the analysis of accuracy rates confirms previous studies (Bartolozzi et al., 2021; Jongman & Meyer, 2017), which found that repetition priming was not modulated by linguistic dual-tasking at encoding.

4.4 Conclusions

In earlier work (Bartolozzi et al., 2021; Jongman & Meyer, 2017), we found that repetition priming was resilient to divisions of attention in a linguistic dual-tasking paradigm. Since previous studies used offline measures (target accuracy rates and naming latencies), it was not possible to determine whether production did not interfere with comprehension at all, or whether primes were processed less thoroughly but still well enough to elicit priming effects. This experiment used an EEG paradigm to assess whether online comprehension processes are affected by concurrent production. The results confirmed that online word processing is rather resilient to division of attention. Although ERPs related to word processing differed in the NP block and the P block, it is not clear whether the effect reflects interference from production to comprehension or simply processing of the to-be-named distractor.

4.5 Appendix A

T.Name	T.Name.Agr	T.Freq	T.length	P.Name	P.Freq	P.length	D.Name	D.Name.Agr	D.Freq	D.length
aansteker	100.00	5.76	1420	schilder	10.50	932	bus	100	64.83	963
aardappel	100.00	3.34	1044	duim	10.79	749	vliegtuig	100	89.92	1140
aardbei	100.00	1.56	1028	zwaan	1.99	836	gezicht	100	183.63	1020
accordeon	84.91	1.19	1165	duim	10.79	749	boerderij	96.49	29.57	917
ananas	100.00	2.52	980	passer	0.11	732	koffie	85.96	133.30	724
anker	100.00	5.03	1237	vrijheidsbeeld	1.72	1357	druif	79.25	1.35	1020
appel	100.00	10.20	717	portemonnee	12.55	1084	duivel	96.55	45.12	980
armband	75.00	6.27	1053	voetbal	12.83	892	schedel	78.57	14.50	989
aubergine	94.34	0.69	884	pet	13.19	604	dak	93.10	54.84	912
avocado	100.00	0.53	1157	badjas	2.26	1077	computer	95	47.89	1092
ballon	100.00	5.28	812	galg	2.52	636	stoel	100	51.20	844
banaan	100.00	5.33	860	trommel	1.85	797	motor	80.95	42.63	940
batterij	100.00	6.75	972	naald	8.51	820	tuinslang	77.97	0.75	940
beer	96.67	25.45	772	schilder	10.50	932	mossel	78.95	0.53	892
beker	78.95	8.78	805	theepot	0.85	949	visser	94.74	4.30	956
bezem	100.00	3.80	1005	gordijn	4.46	980	rimpels	55.56	2.70	1148
bijl	92.86	9.26	796	kruiwagen	1.26	1212	pion	78.95	3.18	804
blad	77.97	11.14	780	munt	10.89	749	dynamiet	51.92	7.23	980
bloemkool	94.83	0.55	1076	robot	12.39	1029	kerkhof	72.88	12.49	949
boog	100.00	8.55	820	schoen	13.45	732	vogel	88.68	32.27	964
bord	82.00	27.30	780	cirkel	12.44	949	jurk	86.21	55.75	789
bril	100.00	24.49	869	trompet	2.68	997	gorilla	77.59	5.12	932
broccoli	94.74	2.65	988	trommel	1.85	797	loodgieter	75.93	5.53	1165
broodrooster	91.07	2.40	1236	luier	3.70	860	hersenen	77.59	35.79	996
brug	87.93	44.07	764	ufo	2.90	820	melk	82.46	39.70	708
cactus	100.00	1.81	965	snavel	1.72	1036	tunnel	100	24.49	764
cadeau	94.83	29.29	820	masker	19.23	797	slot	92.86	52.46	805
citroen	96.55	5.21	940	mand	4.30	717	pop	80.70	23.90	397
dienblad	96.36	1.23	949	oven	9.90	852	bumper	55.36	2.58	989

dobbelsteen	100.00	0.71	1180	schilderij	21.52	1085	mes	100	46.24	780
dolfijn	100.00	1.90	988	scooter	4.89	940	trein	85.96	73.15	797
donut	100.00	4.48	972	brandweerman	5.05	1148	koning	100	138.53	964
douche	89.47	22.25	780	robot	12.39	1029	haar	88.14	2975.87	772
drumstel	100.00	1.51	1036	rok	7.23	644	glas	100	57.15	949
elleboog	100.00	3.27	909	diamant	11.18	989	berg	100	34.30	933
envelop	77.97	6.27	869	paddestoel	0.57	1012	marionet	55.77	2.13	1156
fiets	100.00	21.75	796	brandblusser	1.62	1142	dokter	86.21	244.07	869
gitaar	100.00	11.46	916	schild	9.97	876	kas	92.45	4.37	740
gum	100.00	0.48	740	passer	0.11	732	balkon	78.18	6.63	980
haai	96.49	9.44	717	vleugel	8.83	1004	slaapkamer	84.75	30.83	1276
hamer	100.00	8.55	876	astronaut	4.51	1189	duif	79.66	5.35	796
hek	100.00	22.78	589	kruiwagen	1.26	1212	ring	93.22	52.34	732
helikopter	100.00	21.88	1172	snor	9.95	796	kalender	94.92	3.38	957
helm	93.33	11.09	757	limousine	4.85	1060	foto	85.71	119.16	980
horloge	100.00	28.17	1036	indiaan	7.87	948	skateboard	100	1.83	1260
kaars	100.00	5.88	876	woestijn	27.74	956	rugzak	74.58	7.57	1020
kam	100.00	5.28	637	vierkant	2.63	1117	stopcontact	100	1.28	1357
kameel	80.70	2.70	860	vrijheidsbeeld	1.72	1357	boemerang	100	0.34	1012
kangoeroe	100.00	1.56	980	slager	6.63	1052	hoofd	77.19	274.05	805
kanon	100.00	6.17	797	munt	10.89	749	ambulance	80	25.18	1189
kast	96.49	30.05	732	zonnebloem	0.48	1069	ster	100	43.77	757
kegel	86.27	0.43	877	oorbel	1.69	916	granaatappel	79.59	0.09	1333
ketting	100.00	19.14	812	ufo	2.90	820	zon	100	68.67	749
kip	87.93	37.89	525	theepot	0.85	949	onderbroek	75.93	7.48	1020
kiwi	100.00	0.64	764	galg	2.52	636	rivier	74	52.87	1211
klarinet	61.40	1.03	1052	zwembad	23.17	1052	dief	45.76	29.96	708
knie	100.00	10.24	677	pan	9.38	565	geweer	87.72	55.39	877
knoop	96.61	9.51	1122	zeemeermin	2.79	1165	flipper	75.93	2.24	893
koe	100.00	18.55	524	postzegel	1.67	1053	want	73.33	419.08	789
koekje	73.53	8.12	780	paddestoel	0.57	1012	museum	78.85	19.00	1045
koelkast	100.00	14.80	1029	cirkel	12.44	949	wond	92.59	14.02	772
koffer	96.43	33.87	772	bloem	13.49	684	huis	94.74	818.90	892

kokosnoot	100.00	1.72	1124	woestijn	27.74	956		zwaard	96.43	37.48	1029
komkommer	69.49	1.33	1012	postzegel	1.67	1053		hand	100	199.91	932
kroon	100.00	14.32	764	etui	0.07	765		litteken	85.96	9.81	1060
kruk	96.49	2.56	924	brandweerman	5.05	1148		park	74.07	30.87	756
kurk	95.65	1.62	740	slager	6.63	1052		eikel	96.08	67.03	892
kwast	52.63	1.72	820	masker	19.23	797		pad	72.88	41.99	1472
lepel	100.00	5.01	932	handschoen	6.68	956		put	57.63	11.69	572
liniaal	91.38	0.55	1012	zonnebloem	0.48	1069		bank	94.74	91.91	900
maan	100.00	42.10	717	heks	26.76	740		blouse	78.57	3.68	852
mais	72.88	0.75	949	reuzenrad	1.46	1156		vogelkooi	58.62	0.48	1252
map	77.78	4.53	620	diamant	11.18	989		spuit	73.21	9.79	917
matras	93.55	5.15	996	golf	17.54	836		keuken	100	52.98	900
medaille	96.61	10.18	917	schoen	13.45	732		traktor	91.67	0.07	868
microfoon	100.00	10.34	1140	zwembad	23.17	1052		ontbijt	88.68	44.82	1012
microscoop	96.61	1.92	1164	bijbel	23.05	916		trui	55.56	11.62	732
mier	89.83	2.54	804	vleugel	8.83	1004		puist	76.47	0.96	829
molen	91.38	4.53	932	kraan	6.40	757		lift	92.86	65.47	749
muis	100.00	11.14	852	schilderij	21.52	1085		hengel	73.21	1.94	780
nietmachine	96.67	0.75	1132	caravan	5.12	940		hart	94.92	196.37	732
olifant	100.00	12.01	1116	beeldhouwer	0.85	1157		klaver	38.60	0.50	956
oor	100.00	25.02	724	hamburger	8.60	1084		telefoon	100	156.92	1069
papegaai	100.00	3.29	1029	rok	7.23	644		monster	78.05	49.97	940
paprika	100.00	1.33	989	roos	11.71	900		broekzak	84.48	2.65	949
paraplu	100.00	3.43	909	kraan	6.40	757		stuur	91.38	115.57	884
peer	96.61	1.99	660	limousine	4.85	1060		ridder	76.47	13.58	716
perzik	62.75	2.29	1222	gevangene	28.52	980		boek	100	150.93	612
piano	67.80	14.11	973	spook	12.76	900		druppel	85.96	8.62	836
pijl	100.00	7.11	669	hak	8.19	1083		deur	100	247.48	837
pijp	100.00	13.81	612	schaduw	20.92	997		eiland	100	51.59	1020
pil	93.10	9.40	620	hamburger	8.60	1084		ooglapje	76.36	0.41	1140
pinda	91.23	2.13	829	borstel	2.10	964		herder	75	5.92	852
pizza	100.00	24.38	789	trompet	2.68	997		broek	100	67.28	1201
pompoen	100.00	2.49	876	kasteel	27.60	972		auto	100	458.00	932

pot	84.48	30.62	605	spook	12.76	900	vuur	100	100.57	828
potlood	100.00	5.44	932	handschoen	6.68	956	grot	95	17.45	829
printer	100.00	0.89	884	doolhof	2.63	1084	bom	100	51.32	724
prullenbak	66.10	1.39	932	kasteel	27.60	972	sleutelgat	55.36	0.82	1292
rekenmachine	100.00	0.57	1212	baard	11.64	940	krant	96.43	58.20	724
riem	100.00	14.16	685	puzzel	5.26	725	onderzeeer	77.59	0.73	1132
rits	96.61	4.37	837	borstel	2.10	964	muur	100	66.89	837
saxofoon	83.93	0.98	1220	kaas	22.85	757	lippen	73.21	23.85	853
schaap	100.00	6.54	1234	puzzel	5.26	725	vliegveld	90.74	32.06	1157
schaar	100.00	6.36	740	zwaan	1.99	836	pylon	80.36	0.23	780
schep	75.86	4.62	716	bot	14.52	660	troon	75	12.10	812
schildpad	100.00	4.37	1156	pan	9.38	565	weg	46.30	1481.66	725
schort	94.83	2.81	876	golf	17.54	836	mijn	73.21	4412.02	805
schouder	89.47	18.57	957	brandblusser	1.62	1142	regen	83.93	26.48	965
schroevendraaier	100.00	2.20	1300	apotheek	4.44	1203	walrus	78.95	1.03	1076
sigaret	85.96	27.85	1069	roos	11.71	900	honing	100	7.16	924
skelet	100.00	2.84	1033	voetbal	12.83	892	badkamer	85.96	30.41	1212
sok	100.00	3.11	725	ei	16.19	645	vis	58.33	50.08	820
speen	92.59	0.37	917	indiaan	7.87	948	hond	93.22	168.65	1463
spiegel	100.00	27.44	988	kaas	22.85	757	bos	91.07	46.51	1196
steen	100.00	35.72	892	astronaut	4.51	1189	klauw	86.27	3.27	772
strijkijzer	87.93	0.71	1356	oven	9.90	852	vulkaan	91.53	4.62	1069
stropdas	87.93	4.48	1172	ei	16.19	645	fee	88.68	5.67	725
struisvogel	100.00	0.37	1292	naald	8.51	820	radio	100	58.70	1028
tank	96.55	19.25	876	aquarium	2.88	1157	vrouw	79.25	821.67	789
thermometer	91.89	1.30	1237	pet	13.19	604	jongleur	77.78	0.34	1055
tijger	93.22	11.69	909	beeldhouwer	0.85	1157	kies	54.39	35.35	685
toetsenbord	100.00	1.10	1172	oorbel	1.69	916	parel	85.45	3.02	860
tomaat	100.00	2.97	916	gordijn	4.46	980	rechter	100	63.28	916
t-shirt	54.39	7.64	949	gevangene	28.52	980	snoepje	75.86	6.49	924
ui	87.72	2.33	644	aquarium	2.88	1157	goal	56.60	3.20	865
varken	81.03	24.74	989	etui	0.07	765	badkuip	58.62	3.89	1313
veer	96.67	3.43	860	bot	14.52	660	doos	96.55	38.28	940

ventilator	87.27	2.52	1404	apotheek	4.44	1203	sleutel	100	80.70	1231
veter	92.86	1.33	1020	mand	4.30	717	sla	90.74	83.28	773
vingerafdruk	96.55	4.37	1212	caravan	5.12	940	parachute	100	5.26	1020
viool	75.00	4.30	860	scooter	4.89	940	kompas	100	4.55	1029
vlag	100.00	17.79	612	doolhof	2.63	1084	nacht	83.33	204.44	788
vlieger	100.00	3.80	892	snor	9.95	796	gevangenis	100	104.90	1339
vlinder	100.00	6.13	1029	schaduw	20.92	997	mimespeler	74	0.37	1420
vork	100.00	5.19	820	baard	11.64	940	tamboerijn	79.17	0.43	1069
vuurtoren	100.00	3.36	1292	bijbel	23.05	916	trap	100	52.28	645
walnoot	89.29	0.71	1069	vierkant	2.63	1117	gang	82.46	110.80	724
wandelstok	63.79	0.78	1421	snavel	1.72	1036	voet	91.67	50.81	765
wasknijper	79.31	0.02	1164	zeemeermin	2.79	1165	enkel	92.98	113.68	837
weegschaal	100.00	1.83	1092	hak	8.19	1083	pruik	77.19	6.52	1029
wol	90.74	3.75	677	schild	9.97	876	piloot	89.83	30.12	860
wolk	100.00	5.44	1264	luier	3.70	860	pak	91.23	313.20	557
wortel	96.61	6.11	900	reuzenrad	1.46	1156	parfum	87.72	10.89	868
zaag	100.00	3.54	837	bloem	13.49	684	lade	76.79	2.26	908
zaklamp	100.00	5.08	1308	badjas	2.26	1077	oog	100	68.40	749
zebra	100.00	3.06	924	heks	26.76	740	tafel	87.72	83.40	892
zout	73.21	15.50	844	portemonnee	12.55	1084	kaarten	56.36	39.65	972

Table 4.5: Name agreement, frequency, and length of audio recordings of unrelated primes, identical primes/targets, and distractors, when applicable. T = Target, P = Prime, D = Distractor, D.length = length of distractor sound file in milliseconds.

4.6 Appendix B

In this section, we report the results of the behavioural analyses that only included the participant datasets used in the EEG analysis (n=30).

Study phase

	Distractor accuracy rate and mean naming latency		
Relatedness	Accuracy rate	Mean Naming latency (ms)	SD (ms)
Identical	0.72	429	175
Unrelated	0.77	418	166

Table 4.6: The table includes distractor accuracy rates and naming latencies (mean and standard deviation) by relatedness of the 30 datasets used in the EEG analysis.

In the analysis of accuracy rates, the Likelihood Ratio Test showed that adding relatedness as a predictor did not improve model fit ($\chi^2(1) = 3.699$, $p = .054$). The same held in the analysis of distractor naming latencies ($\chi^2(1) = 1.12$, $p =$ .290)

Test phase

		Target accuracy rate and mean target naming latency		
Relatedness	Block	Accuracy rate	Mean Naming latency (ms)	SD (ms)
Identical	*Plan*	0.90	449	171
Identical	*No Plan*	0.92	485	195
Unrelated	*Plan*	0.88	459	186
Unrelated	*No Plan*	0.88	493	199

Table 4.7: The table includes target accuracy rates and naming latencies (mean and standard deviation) by relatedness and block of the 30 datasets used in the EEG analysis.

In the analysis of target accuracy rates, neither relatedness ($\chi^2(1) = 1.85$, p = .174), block ($\chi^2(1) = 2.62$, $p = .105$) or their interaction ($\chi^2(1) = 2.59$, p = .108) improved model fit. In the analysis of target naming latencies, adding block as a predictor improved model fit ($\chi^2(1) = 4.68$, $p = .031$), while adding relatedness ($\chi^2(1) = 0.77$, $p = .379$) and the interaction did not ($\chi^2(1) = 0.3$, $p = .587$).

Fixed effects				Random effects		
Fixed	Estimate	SE	t	Random	Variance	sd
Intercept	-0.004	0.05	-0.09	Target picture (Intercept)	0.013	0.11
Block	-0.029	0.01	-2.25	Target Picture (Relatedness)	0.001	0.04
				Participant (Intercept)	0.061	0.25
				Participant (Block)	0.004	0.06

Table 4.8: The table includes the model with (log-transformed) picture-naming latencies as the dependent variable and Block as the independent variable. This model was run with n=30 (same participants included in EEG analysis).

4.7 Appendix C

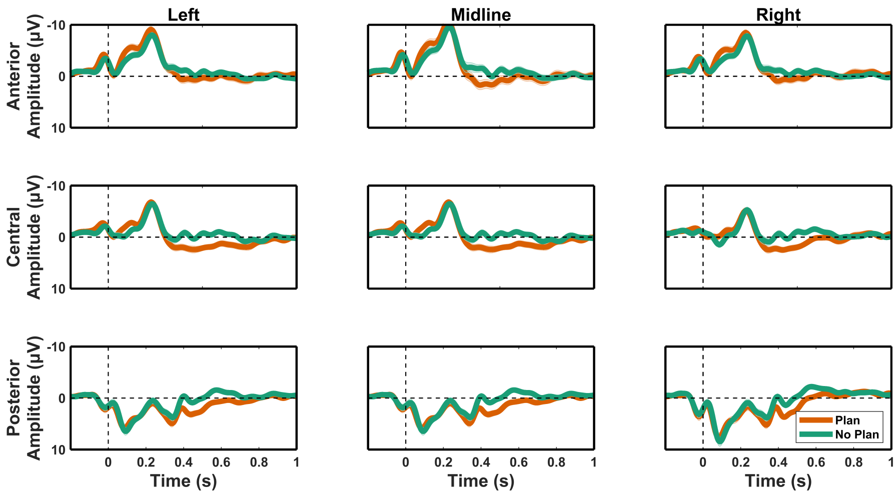

Figure 4.7: Overview of ERP waveforms in the study phase (NP vs P blocks) for electrodes in left anterior, mid anterior, right anterior, left central, mid central, right central, left posterior, mid posterior, and right posterior regions.

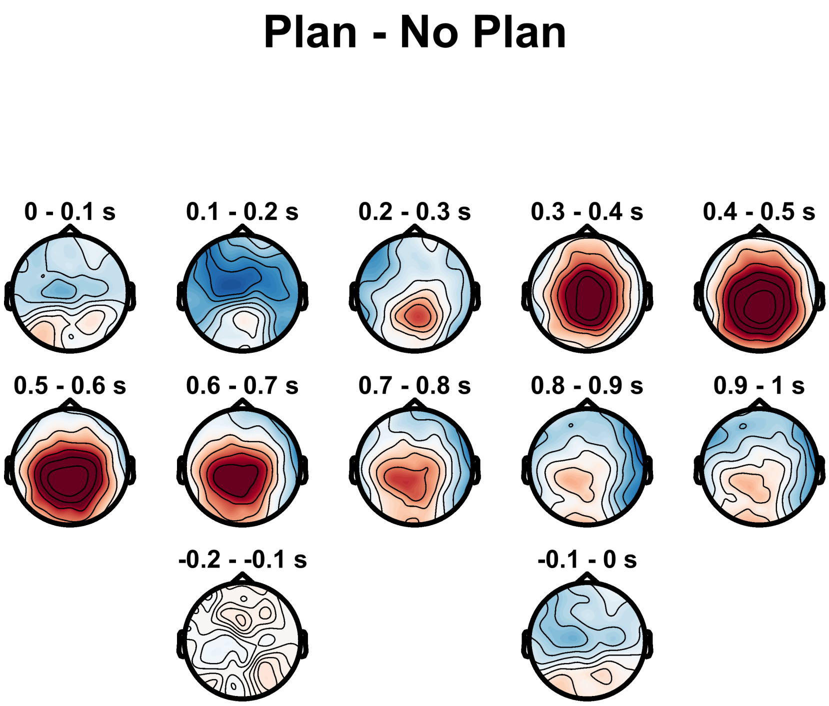

Figure 4.8: Overview of scalp distributions in the study phase (NP vs P blocks) from 0ms to 1s after stimulus onset. The topographies in the bottom row describe activity in the time window before prime word onset (distractor picture onset: -150ms).

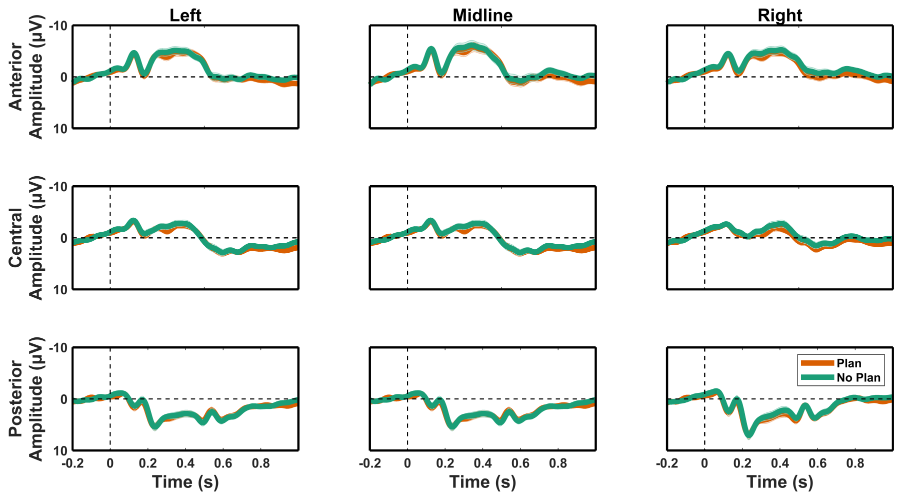

Figure 4.9: Overview of waveforms in the test phase (NP vs P blocks) for electrodes in left anterior, mid anterior, right anterior, left central, mid central, right central, left posterior, mid posterior, and right posterior regions.

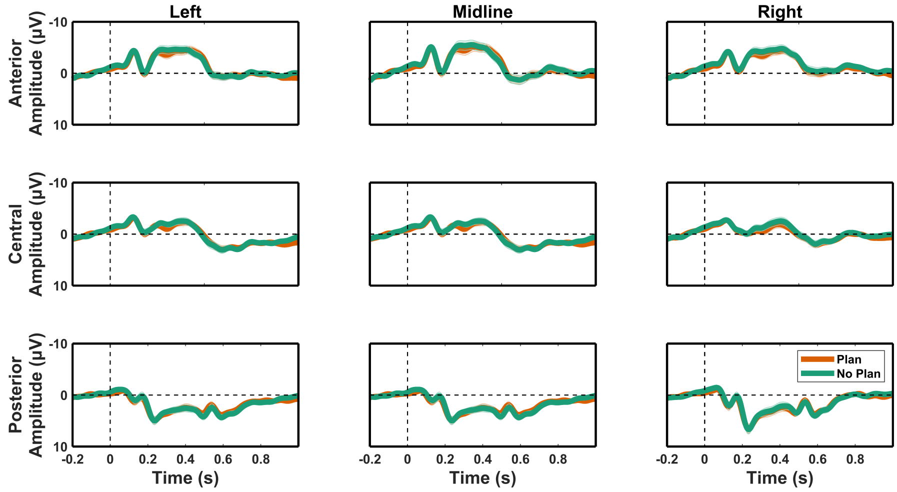

Figure 4.10: Overview of waveforms in the test phase (NP vs P blocks) after excluding three participants for electrodes in left anterior, mid anterior, right anterior, left central, mid central, right central, left posterior, mid posterior, and right posterior regions.

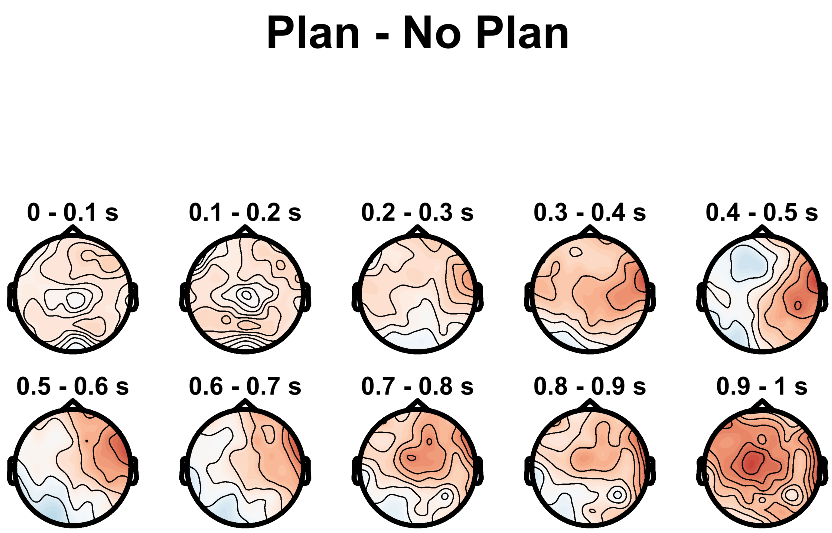

Figure 4.11: Overview of scalp distributions in the test phase (NP vs P blocks) from 0ms to 1s after stimulus onset.

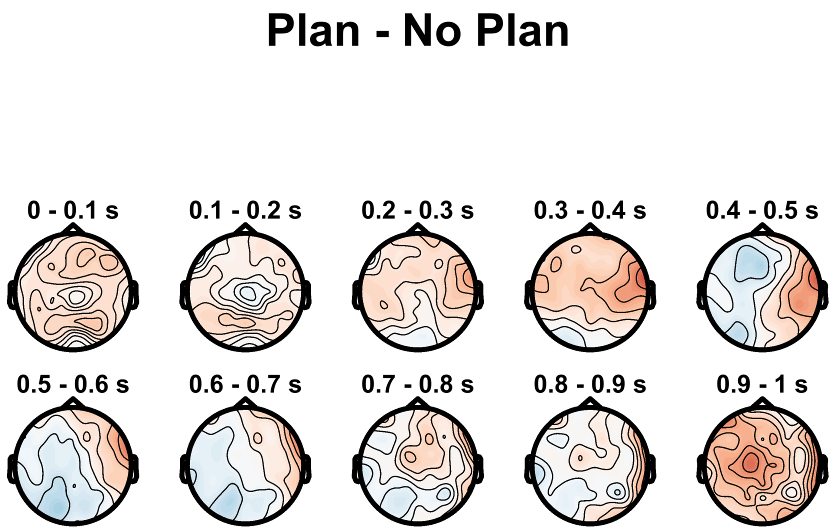

Figure 4.12: Overview of scalp distributions in the test phase (NP vs P blocks) from 0ms to 1s post stimulus onset after excluding three participants.

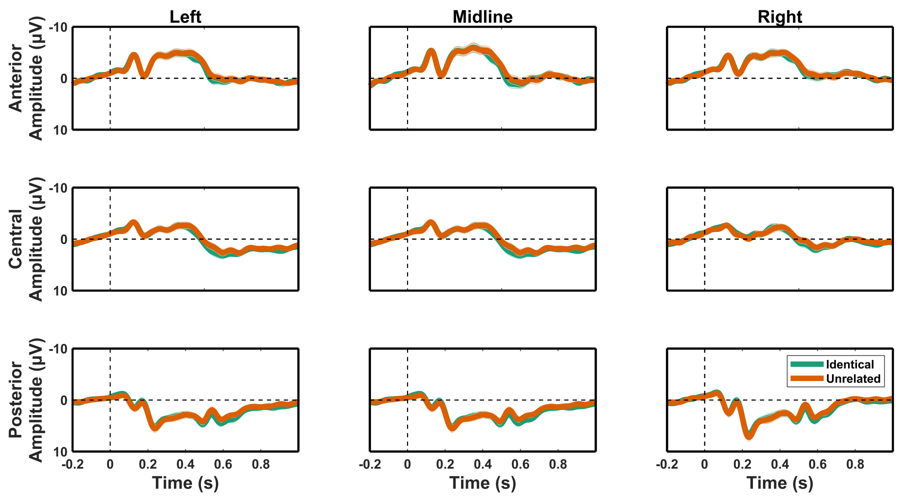

Figure 4.13: Overview of waveforms in the test phase (identical vs unrelated trials) for electrodes in left anterior, mid anterior, right anterior, left central, mid central, right central, left posterior, mid posterior, and right posterior regions.

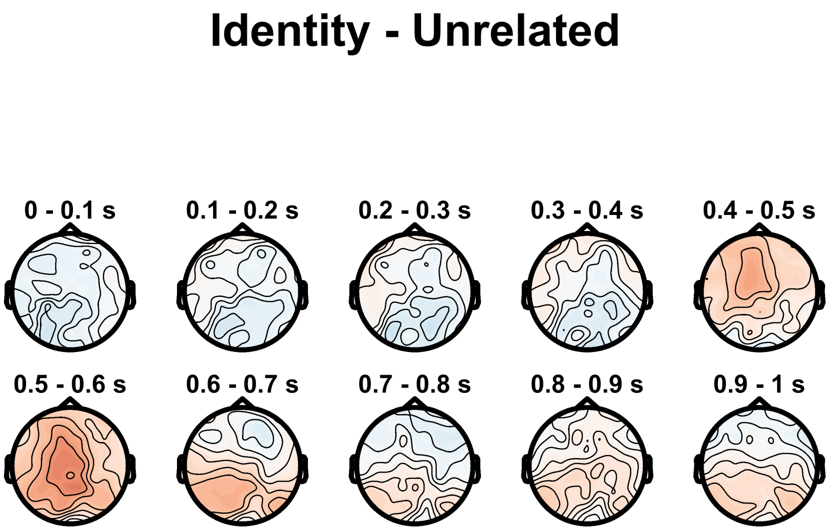

Figure 4.14: Overview of topographies in the test phase (identical vs unrelated trials) from 0ms to 1s post stimulus onset.

5 Repetition priming in comprehension: evidence from EEG

Abstract

Repetition priming has been shown to be attenuated or eliminated when the prime word is presented as part of a sentence. In this study, we analysed a publicly available dataset to determine whether repetition priming effects can be observed during naturalistic comprehension of a story. Using a linear deconvolution approach, we modelled the EEG signal against a predictor indicating whether a word had occurred earlier in the text. We predicted that repeated words would elicit a reduced N400 with respect to non-repeated words. While a first model confirmed our hypothesis, repetition priming effects disappeared when the effect of word predictability was also taken into account in the model. We suggest that our pattern of results arose because predictability effects are stronger than priming effects in naturalistic comprehension but we also discuss factors that may have undermined our ability to detect priming in the current study.

Keywords: N400, repetition priming, predictability, naturalistic comprehension

5.1 Introduction

Repeating a part of an utterance is associated with faster and/or more accurate processing. This phenomenon, known as priming, has been widely investigated since the second half of the last century (Bargh, 2014). In the last twenty years, psycholinguistic studies have focused on priming as a possible supporting mechanism in conversation, as posited by an influential model of dialogue, the interactive alignment account (Pickering & Garrod, 2004, 2013). Repetition benefits have been shown at different levels of processing, such as the syntactic, conceptual, word-form, phonological (e.g., Hamburger & Slowiaczek, 1996; Jongman & Meyer, 2017; Pickering & Branigan, 1999; Wheeldon & Monsell, 1992). Notwithstanding the number of studies investigating priming, less is known about priming in naturalistic settings. In this study, we set out to answer this question by investigating correlates of priming in the EEG signal of participants listening to a story, with no other task demands. We specifically focused on the repetition of words (repetition priming), which has been widely investigated in the psycholinguistic literature and has been shown to yield a reliable and strong effect (e.g., Forster & Davis, 1984; Wheeldon & Monsell, 1992). Before turning to the details of the experiment, we describe previous studies of repetition priming and the factors that have been shown to modulate the magnitude of the effect.

5.1.1 Repetition priming and the factors modulating the effect

Repetition priming effects have been shown in different modalities (visual, auditory, cross-modal; e.g., Besken & Mulligan, 2010; Durso & Johnson, 1979; Nicolas & Söderlund, 2000) and using a variety of techniques (e.g., lexical decision task, picture naming, perceptual identification task; see Jacoby (1983); Jongman and Meyer (2017); Oliphant (1983)). In behavioural experiments, repeated words usually elicit faster reaction times or higher accuracy rates compared to non-repeated words (e.g., MacLeod, 1989; Wheeldon & Monsell, 1992). Electrophysiological studies have shown that repetition priming is associated with modulations of the N400, i.e., a negative-going component with a centroparietal distribution peaking around 400ms post stimulus onset (Kutas & Federmeier, 2011). In particular, repeated words elicit a reduced N400 relative to non-repeated words (Besson, Kutas, & Petten, 1992; Kutas & Federmeier, 2011; Rugg, 1985).

One consistent finding across repetition priming studies is that not all words are primed to the same degree and that the effect can be modulated by a variety of factors. An important factor affecting repetition priming is whether a prime word (i.e., the first occurrence of a word) occurs alone or in an embedding at first presentation. Words in a list elicit more priming of later targets than words presented as part of a sentence (e.g., Besken & Mulligan, 2010; Masson & Macleod, 2000). For instance, over the course of five experiments, Besken and Mulligan (2010) measured repetition benefits of prime words that had been presented in isolation, as part of meaningful sentences, or as part of scrambled sentences. Priming, measured using a stem or fragment completion task, was greater for prime words in lists than from prime words in sentences and comparable for prime words in meaningful and scrambled sentences.

While Besken and Mulligan (2010) found that embedding a prime word in a sentence decreased priming, regardless of its meaningfulness, further studies have shown that the content of the sentence in which the prime word is embedded can also modulate the effect (Hodapp & Rabovsky, 2021; Rommers & Federmeier, 2018a, 2018b). For instance, in an EEG experiment by Rommers and Federmeier (2018b), participants read series of four sentences. The prime word in the first sentence was preceded by a weakly or highly constraining context. After two filler sentences, the target word was presented in the fourth sentence, this time preceded by a new weakly constraining context. The authors hypothesised that the initial processing of prime words would be enhanced when they were preceded by a weakly - rather than highly - constraining context, therefore yielding more priming at subsequent presentation. Indeed, critical words initially presented in weakly constraining contexts elicited a greater N400 reduction at repetition with respect to words initially presented in highly constraining contexts. To explain these results, the authors suggested that words preceded by highly constraining sentences were not processed as thoroughly as words preceded by low-constraining sentences. Such processing affects the quality of linguistic representations created at first presentation and, in turn, priming effects at subsequent repetition.

In another EEG experiment, Hodapp and Rabovsky (2021) asked participants to read sentences that contained either expected or unexpected (but plausible) words. After a filler working memory task, participants carried out a perceptual identification task, where masked words were shown with progressively increasing display time. As predicted, the perceptual identification times were shorter for previously unexpected words than previously expected words. A first linear mixed effects model confirmed that log-transformed perceptual identification times were modulated by previous expectancy (expected vs non-expected), repetition (repeated vs previously not seen) and frequency. In a subsequent model, log-transformed perceptual identification times were modelled against the amplitude of the N400 during sentence reading and word frequency. As in the previous analysis, likelihood ratio tests confirmed the significance of word frequency. While the modulation of the N400 amplitude on the perceptual identification times failed to reach statistical significance, the RT difference between previously expected and unexpected items correlated with the corresponding N400 difference in the reading task at the participant level. The authors argued that the nonsignificant effect of the N400 amplitude on the reaction times depended on the fact that the reaction times and the N400 amplitude elicited by high-frequency words are inherently smaller than those elicited by low-frequency words. As a result, any relationships between perceptual identification times and the N400 amplitude only emerge when the effect of frequency is controlled for.

In sum, the studies described above suggest that embedding a word in a sentence affects the magnitude of priming effects and that the type of embedding (e.g., predictable vs non predictable) may modulate the effect. In most of these studies priming was still evident (but reduced) when words were embedded in sentences. However, whether repetition effects also occur during naturalistic comprehension of stories is an empirical question.

For instance, one main difference between lab experiments and naturalistic comprehension concerns the context in which the repeated word was presented: in the studies above (except for Rommers and Federmeier (2018b)), priming was always tested by presenting words in isolation, e.g., in a fragment completion task or in a perceptual identification task. Such tasks differ from comprehension in naturalistic settings, in which words are embedded in sentences at first and subsequent presentations. If the type of embedding affects the way the word is processed, this should be true both for words at first and subsequent presentation. It is therefore possible that priming effects might be attenuated when the repeated word is also embedded in a sentence. Furthermore, in the lab experiments described in this section, the embeddings of the words at first presentation only consisted of one or two short sentences. In longer texts, a variety of other factors - such as the number and type of intervening words, and the number of repetitions of the words, their sentential positions - may affect word processing, and in turn affect priming.

5.2 The current study

Overall, the studies mentioned above suggest that repetition priming is modulated by the embedding of a word when it is first presented. Since specific characteristics of the embedding might affect repetition benefits for words, it is important to determine whether repetition priming also occurs in naturalistic settings, where not only the embedding of a word but also other factors may contribute to mitigate the effect. Investigating this question is relevant, especially in light of models that view priming as a pivotal mechanism in conversation (Pickering & Garrod, 2004, 2013). While in this study we focused on comprehension, determining the occurrence and possible limitations of the effect is relevant for theories focusing on how language processing unfolds in natural settings more generally.

We asked two main questions. First, we asked whether repetition priming effects are present when participants listen to a story, with no other task demands. Second, we asked whether repetition priming effects were affected by the predictability of a word at first presentation, as suggested by Rommers and Federmeier (2018b) and Hodapp and Rabovsky (2021). In order to answer these questions, we reanalysed a publicly available EEG dataset collected while participants were listening to part of an audio book (Broderick, Anderson, Di Liberto, Crosse, & Lalor, 2018; Di Liberto, Wong, Melnik, & de Cheveigné, 2019).

5.2.1 Materials

The EEG data used in the analyses belong to a publicly available dataset recorded while participants listened to part of a narrative story, which was split in 20 runs, each lasting around 180s and containing about 600 words (Broderick et al., 2018). The EEG dataset contained information about 5459 word tokens, of which 1508 were unique words (see Table 5.2 in Appendix A for the distribution of words in different lexical categories). Of the 1508 words, 706 were repeated once or multiple times across the dataset. The first and last word of each run were excluded from the analyses.

For each word, we derived frequency values from the SUBTLEX-UK database (Van Heuven, Mandera, Keuleers, & Brysbaert, 2014). The database reports frequency values on a logarithmic scale, ranging from 1 (very low frequency) to 7 (very high frequency). Word predictability was quantified using lexical surprise values. The lexical surprise values ($\log(p(word \mid context))$, where p = probability) used in this experiment were obtained from a previous study, that used a pre-trained language model, GPT-2, on the same dataset (Radford et al., 2019, see Heilbron, Armeni, Schoffelen, Hagoort, and de Lange (2021) for details on how the surprise values were derived). Heilbron et al. (2021) calculated surprise values for each run separately: when calculating the surprise value for a word, the whole run up to word x was taken into account. Information concerning the frequency and surprise distribution of the word tokens in the dataset can be found in Figures 5.3, 5.4, and 5.5 of Appendix A.

For the main analyses, we categorised word tokens as first or only occurrence tokens versus repeated tokens (across all blocks). The former category included all tokens in the corpus. The latter category included a subset, namely those words that were repeated at least once. There were 3951 word tokens (out of 5459 tokens) in this subset. Words could be repeated with different lags, ranging from 1 word (immediate repetition) to 5183 words. The distribution of distances between successive repetitions of a word (e.g., from first occurrence to first repetition, from first repetition to second repetition etc., calculated across all runs in the dataset) is shown in Figure 5.6 of Appendix A.

The main analyses compared the EEG signal for first or only tokens (i.e., the whole set) against the signal for the repeated subset. Repetitions were always calculated across runs. In Appendix C we report post-hoc exploratory analyses comparing the EEG signal for first instances and second instances (i.e., first repetitions) of a word (models 1A-4A, 6A in Appendix C). We also report analyses where repetitions were calculated within runs (models 5A-6A in Appendix C).

	mean	sd	min	max
frequency per word	4.36	1.02	1.17	6.55
number of repetitions per word	3.62	8.02	1.00	129.00
intervening items between repetitions	426.97	709.77	1.00	5183.00
word surprise values	4.67	3.07	0.00	20.38

Table 5.1: The table provides descriptive information concerning frequency per word (logscale from SUBTLEX-UK), number of occurrences per words (1 = word occurs only once in the text), number of intervening words between successive repetitions (calculated across all items in dataset; 1 = words follow each other) and surprise values across all items in the dataset.

5.2.2 Planned analyses

The first two planned analyses were designed to answer our first research question, i.e., whether repetition priming occurred during comprehension of a naturalistic text. In the first analysis, we planned to regress the EEG signal against a predictor for repetition (first or only occurrence of word vs repeated word), a predictor for word frequency, and a predictor for the interaction between word frequency and repetition. Repetition was a binary predictor indicating whether the word was the first or only occurrence in the text (hereafter, non-repeated word) or whether it was a repetition of a word that had occurred earlier in the text (hereafter, repeated word). This binary predictor did not take into account whether a word was the first or nth repetition. Concerning the effect of repetition, we hypothesised that both repeated and non-repeated words would elicit a negative-going response between 300ms and 500ms post word onset, compatible with a N400. We hypothesised that words classified as repeated would elicit a reduced N400 with respect to non-repeated words. Word frequency and the interaction between frequency and repetition were added to the model because previous priming studies using words embedded in sentences, and presenting them auditorily, showed priming for low-frequency but not high-frequency words (MacLeod, 1989; Nicolas & Söderlund, 2000). We therefore argued that, if embedding words in sentences diminished the magnitude of priming, repetition benefits should still be evident at least for low-frequency words. Given that the N400 is modulated not only by repetition but also by predictability (Kutas & Federmeier, 2011), we planned a second model, where we regressed the EEG signal against the same predictors as in the first analysis, this time also taking into account the effect of the predictability of a word upon presentation (hereafter, current predictability).

The last planned analysis focused on our second research question, i.e., whether repetition priming effects were affected by the predictability of the word at first presentation (hereafter, previous predictability). We set out to answer this question in a model with the EEG signal as the independent variable, previous predictability as the predictor of interest, and word frequency as a covariate. We hypothesised that repetition priming would be greater for words with low rather than high previous predictability. This analysis was contingent to the first two: indeed, it could only be implemented if the previous two models showed reliable priming effects. To anticipate the results of this study, we could not run this model, due to the instability of priming in the preceding models.

5.2.3 Participants

19 native speakers of English (13 male) took part in the experiment (age range: 19-38).

5.2.4 Data preprocessing

We ran EEG preprocessing steps in MATLAB (R2020b; Mathworks, Inc.) using the Fieldtrip toolbox (Oostenveld et al., 2011). We first high-pass filtered the data at 1Hz and low-pass filtered the data at 60Hz, using a windowed sinc finite impulse response filter. We then re-referenced the data to the average of electrodes placed on the left and right mastoids. Preprocessing and artifact rejection were performed on the combined data from all runs for each participant dataset.

Artifacts in the EEG data were dealt with in three steps: first, we inspected the data visually, removed channels containing extreme values, and marked portions of the data that also contained extreme values (all time segments containing artifacts were kept in the data, but set to zero in the design matrices when running regression models). Next, we performed ICA (*runica* implementation in Fieldtrip), identified components related to eye movements and eye blinks (average number of components across participants: 2.21; range: 1-5), and recombined the remaining components. We then performed a final visual inspection of the data and marked any remaining artifacts. Bad electrodes that were removed in the first step were then recovered from the mean activity of a combination of surrounding electrodes.

Data were low-pass filtered again at 8Hz, in agreement with previous studies that employed the same dataset (Broderick et al., 2018; Broderick, Anderson, & Lalor, 2019; Heilbron et al., 2021, 2019). We then re-referenced to the average of the left and right mastoids. The data were then converted from MATLAB to EEGLAB-compatible format, i.e., the default format of the Unfold toolbox which was used for regression modelling. Before running the analyses, we z-scored the EEG data, and centered and scaled all predictors.

5.2.5 Data analysis

In naturalistic settings, the EEG activity at a given time point can be considered as the summed response to multiple, partially overlapping stimuli. For instance, in the data used in this experiment, the EEG response at a given time point corresponds to the activity elicited by multiple words occurring in sequence, and whose associated neural response therefore also overlaps in time. This can lead to difficulty in interpreting neural responses to events (words) of interest in the data, because those responses are contaminated with activity related to nearby events. One way to circumvent this issue is to use regression-based deconvolution approaches. Starting from the recorded signal, one can model the onsets of the events that are thought to elicit the signal (e.g., each word presented to the participant) and a well-specified model will return the isolated neural response associated with each event (Ehinger & Dimigen, 2019). In other words, given the EEG signal - i.e., the independent variable - the model will return the betas (or regression-ERPs) of the predictors of interest. For this reason, the data were analysed using the Unfold toolbox (Ehinger & Dimigen, 2019), which provides tools for regression-based EEG analysis and overlap correction.

Since the EEG signal indexes the summed activity of multiple events, the model matrix needs to be time-expanded before running the regression (Dimigen & Ehinger, 2020; Ehinger & Dimigen, 2019). The model matrix before time-expansion contains as many rows as the number of events in the data, and as many columns as the number of predictors (and the intercept). After the time expansion, the matrix contains as many rows as the number of time samples in the data and as many columns as the number of predictors, corresponding to the time lag chosen (in this study, from -0.2s to 0.9s). In other words, for each predictor, the time-expansion adds as many columns as the number of samples in the time window, meaning that each column represents the predictor at a specific time lag with respect to the onset of the event.

The linear deconvolution models in this study were run using the default LSMR solver (Fong & Saunders, 2011): each model simultaneously modelled the relationship between the EEG signal at each word and the predictors of interest (including word onset) separately for each electrode and each participant across all runs in the dataset. The resulting coefficients of the predictors of interest were then compared against zero using threshold-free cluster enhancement (TFCE, S. M. Smith & Nichols, 2009) to probe whether, and where (in space and time) a particular predictor of interest's marginal contribution to the model was statistically reliable across participants.

In a first model, in addition to word onsets, we regressed the EEG signal against repetition (a binary predictor indicating whether or not the word was an instance of repetition within the story), word frequency, and their interaction. In a second analysis, we added lexical surprise as a covariate to the first model to determine whether repetition effects are still evident when the predictability of the word is taken into account. These two models were run to determine whether repeated words elicited an attenuated N400 with respect to non-repeated words.

A third model planned to regress the EEG signal of repeated words against word frequency, and the predictability of the first instance of any repeated word (quantified by the surprise value at the previous instance of the repeated word, i.e., previous surprise). The aim of this model was to assess if predictability of a word at first presentation influenced subsequent priming effects. Whether or not it made sense to run the third model was contingent on finding a reliable repetition effect, and so this third analysis was not carried out.

Before running our models of interest, we ran a baseline model (model 0) to validate that our modelling approach captured known modulations of the N400 by word predictability. To do so, we regressed the EEG signal against the GPT-2 surprise values, and added word frequency as a covariate. Since many previous studies showed a relationship between predictability and the N400 (see Kutas and Federmeier (2011) for a review and Heilbron et al. (2019) for a study using this dataset), we expected that more predictable words (i.e., with low lexical surprise values) would be associated with a reduced N400 with respect to less predictable words (i.e., words with high surprise values). All models are summarised below:

Model 0

$$EEG \sim 1, EEG \sim frequency + surprise_{current} \tag{5.1}$$

Model 1

$$EEG \sim 1, EEG \sim cat(repetition) + frequency + cat(repetition) : frequency \tag{5.2}$$

Model 2

$$\mathrm{EEG} \sim 1, EEG \sim cat(repetition) + frequency + cat(repetition) : frequency + surprise_{current} \tag{5.3}$$

Model 3

$$EEG_{repeated} \sim 1, EEG_{repeated} \sim frequency + surprise_{previous} \tag{5.4}$$

5.2.6 Results

Model 0. We ran a TFCE analysis to test the effect of the predictor for surprise, which was statistically significant ($p < .05$). On the basis of the TFCE output (see Figure 5.1), the effect is likely to be driven by three main clusters. We identified a first early negative cluster, spanning from around 100ms to 250ms post word onset, and with a parietal distribution. This showed that words with high surprise values were associated with a more negative-going ERP than words with low surprise values. The main cluster of interest for the purposes of the analysis is one ranging between around 250ms and 500ms post word onset. The timecourse and the centro-parietal distribution of the effect are consistent with a N400 component (Kutas & Federmeier, 2011). In particular, the data showed that the higher the surprise values were (i.e., the less predictable the word was) the more negative was the N400 amplitude. In addition to the two negative clusters, we also identified a positive cluster between around 300ms and 500ms.

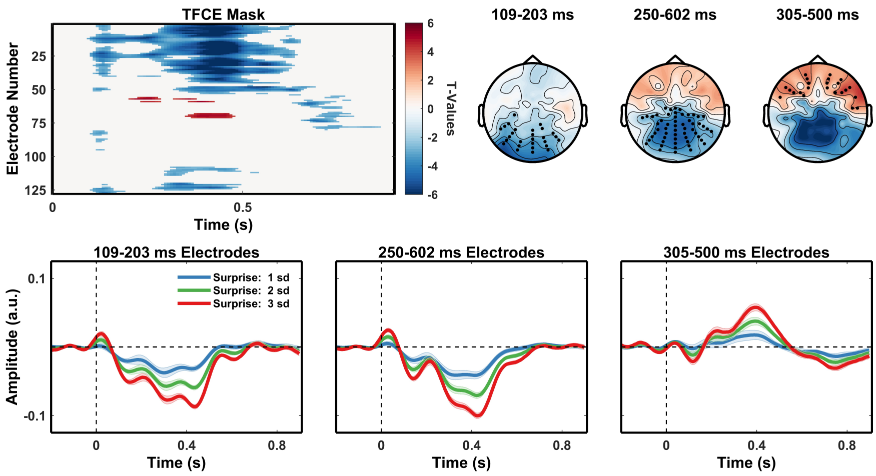

Figure 5.1: Effect of surprise in model 0. Top left: significant clusters in the TFCE analysis. *Top right*: scalp topographies showing the difference between the averaged coefficients of the predictor for surprise and zero. The black marks indicate electrodes contributing to each cluster (significant at alpha < 0.05, two-tailed). *Bottom*: in the model, the intercept captures the effect when the surprise predictor is 0. The waveforms depict the timecourse of the sum of the intercept and the marginal coefficients for the surprise predictor at increasing standard deviation from the predictor mean (1SD, 2SD, 3SD). The waveforms were averaged over the electrodes contributing to each cluster (standard errors of the mean indicated by shaded area around the lineplots).

Model 1 and Model 2. In model one, we regressed the EEG signal against repetition and word frequency, and then used the TFCE test to compare the coefficients of repetition and of the interaction between repetition and frequency against zero. The effect of repetition was significant, while that of the interaction was not. Concerning the coefficients of repetition, the TFCE yielded a positive cluster between around 360 and 460ms, with a centro-parietal distribution ($p <$ .05). The distribution, polarity and time course of the effect are consistent with the N400 component. In particular, the line plot (Figure 5.2, bottom left) shows that repeated words led to a reduction in amplitude of the N400 with respect to non-repeated words. This outcome is in line with previous EEG studies, according to which word repetition leads to the attenuation of the N400 (Besson et al., 1992; Kutas & Federmeier, 2011; Rugg, 1985).

In a second analysis, we investigated whether repetition effects were still present when controlling for the influence of current predictability by adding current surprise values to the model as a covariate. As in the case of model 1, we ran a TFCE test to determine whether the coefficients of the predictors for repetition and for the interaction between repetition and frequency were different from zero. While the plot describing the topography and time course of the predictor for repetition shows a short-lived difference between repeated and non-repeated words (see Figure 5.10 in appendix B), the TFCE analysis did not yield any clusters ($p > .05$) for repetition, showing that the effect was not statistically reliable. As in model 1, the effect of the interaction was not statistically significant.

These results suggest that, when one controls for current predictability, any repetition effects are overshadowed by the relationship between the N400 and predictability. To confirm that this was the case we plotted the topography and timecourse of the surprise covariate from model 2 (Appendix B, Figure 5.11), and it does indeed appear to show a strong relationship with the amplitude of the N400. Given that our repetition priming effects turned out to be unreliable when controlling for word predictability, suggesting that repetition may have a limited role in naturalistic comprehension of texts, we did not carry out the final contingent planned analysis.

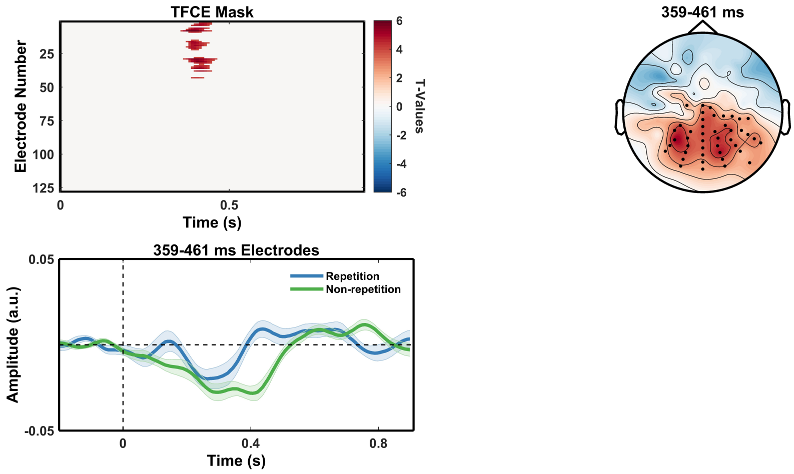

Figure 5.2: Effect of repetition in model 1. Top left: significant cluster in the TFCE analysis. *Top right*: scalp topographies showing the difference between the averaged coefficients of the predictor for repetition and zero. The black marks indicate electrodes contributing to each cluster (significant at alpha < 0.05, two-tailed). *Bottom*: in the model, the intercept captures the effect when there is no effect of repetition. The waveforms depict the timecourse of the intercept (non-repetition) and the sum of the intercept and the coefficients of repetition. The waveforms were averaged over the electrodes contributing to the cluster (standard errors of the mean indicated by shaded area around the lineplots).

5.3 Discussion

In this study, we analysed a publicly available EEG dataset of participants listening to an audio book to answer two main questions (Broderick et al., 2018; Di Liberto et al., 2019). The first research question was whether and how repetition priming effects unfold during naturalistic comprehension. Answering this question is important because it can contribute to theories of dialogue, which assign a pivotal role to priming as supporting mechanism in speaking and listening (Pickering & Garrod, 2004, 2013). The second research question was whether repetition priming effects are modulated by the embedding of the word at first presentation (i.e., previous predictability). This question was motivated by the fact that such modulations have been found in two lab experiments (Hodapp & Rabovsky, 2021; Rommers & Federmeier, 2018b) but no study - that we know of - has focused on whether they also occur during naturalistic comprehension of texts.

In particular, we formulated two main hypotheses. Given previous electrophysiological studies on repetition priming (Besson et al., 1992; Kutas & Federmeier, 2011; Rugg, 1985), we hypothesised that repetition would modulate the amplitude of the N400. Concerning our second question, we predicted that repeated words with high surprise values at first presentation would show a more reduced N400 at repetition than repeated words with low surprise values at first presentation. In order to answer these questions, we planned to run three models where we regressed the EEG signal against our predictors of interest. These main analyses were preceded by a baseline model, where we regressed the EEG signal against surprise values and against frequency (covariate). We ran this baseline model to ascertain that our modelling approach was able to detect modulations of the N400. Indeed, previous studies highlighted a strong association between predictability and the N400 (Kutas & Federmeier, 2011), which was also confirmed in a previous study using our same dataset (Heilbron et al., 2019).

As described in the results section, we were not able to test the relationship between previous word predictability and subsequent repetition effects. For this reason, the following paragraphs will briefly describe the results of the baseline model, and then focus on possible explanations concerning the lack of reliable priming effects, outlining suggestions as to how future experiments could overcome the limitations of this study.

5.3.1 Predictability effects in the baseline model

The results of the baseline analysis showed a significant modulation of the EEG signal by lexical surprise. We suggested that the effect was driven by three main clusters. The main cluster of interest for the purposes of the analysis is one ranging between around 250ms and 500ms post word onset. The timecourse and the centro-parietal distribution of the effect are consistent with a N400 component (Kutas & Federmeier, 2011). In particular, the data showed that the higher the surprise values were (i.e., the less predictable the word was) the more negative was the N400 amplitude (see Figure 5.1). This finding is consistent with previous findings that word predictability is inversely related to the amplitude of the N400 component.

This N400 effect was preceded by an earlier negative cluster, spanning from around 100ms to 250ms post word onset, and with a parietal distribution. As in the case of the N400-like effect, the second cluster showed that words with high surprise values were associated with a more negative-going ERP than words with low surprise values. This early negative component has been found in previous studies investigating predictability effects in auditory comprehension (see

Nieuwland (2019) for a review). For instance, in a study by Hagoort and Brown (2000), sentences containing semantic violations were associated with a negative deflection in a time window between 200ms and 300ms post stimulus onset.

In addition to the two negative clusters, the TFCE also yielded a positive cluster between around 300ms and 500ms. This positivity is consistent with a P3a, an ERP with a centro-frontal distribution and with a timecourse consistent to the one observed in this study (Polich, 2003, 2007). A P3a, also known as novelty P3a (Friedman, Cycowicz, & Gaeta, 2001) can be elicited when an infrequent distractor stimulus is presented among a series of frequent stimuli and it is hypothesised to index allocation of attention to the stimulus. In this study, the positivity was more pronounced for less predictable words than for more predictable words. While we are not aware of studies that found a modulation of linguistic predictability on the amplitude of the P3a, it is possible that the occurrence of less predictable words - as in the case of infrequent stimuli - may be associated with an increase of resources allocated to the stimulus for further processing. However, given that we did not formulate any predictions concerning this component, we are cautious in the interpretation of the effect and will not discuss it further.

5.3.2 Instability of priming effects and suggestions for further research

The combined outcomes of the first and second models showed that priming effects in this study were not reliable: while repetition modulated the amplitude of the N400 in the first model, the effect disappeared when a predictor for surprise was added as a covariate. Although the first model highlighted a modulation of the EEG signal by repetition, this effect vanished when controlling for the predictability of the repeated/non-repeated word. Indeed, the effect appears to be driven entirely by the predictability of the repeated word. Below, we outline some hypotheses about the lack of repetition priming effects in this study.

While a series of studies (Rommers & Federmeier, 2018a, 2018b) showed that predictability of a sentence at first presentation modulates repetition priming at subsequent repetition, in these experiments target words were always embedded in weakly constraining contexts that did not afford any predictions. Instead, in this study the predictability of the sentences could vary both at first and at subsequent repetitions. Based on our results, it is therefore possible that predictability effects are more pervasive than priming during naturalistic comprehension and are sufficient to guide processing. According to this interpretation, local context may be more relevant for processing than across-story repetition.

This does not necessarily mean that priming does not occur in naturalistic comprehension but suggests that its role may be limited to specific circumstances. For instance, priming may be beneficial when predictability fails to aid processing, such as when participants hear an unexpected or incongruent word in the text, in line with accounts that posit priming as the result of an error-based mechanism (e.g., Howard et al., 2006; Oppenheim et al., 2010; Rabovsky et al., 2018). In this case, more attentional resources may be allocated to this word, so that its processing is facilitated at subsequent repetition. In narrative texts such as the one used in this experiment, incongruent or highly unexpected words are rare (see Figure 5.4 and 5.5 in Appendix A for the distribution of surprise and frequency values in the EEG datset), therefore affecting our ability to detect priming effects. In addition, it must also be pointed out that, according to models of conversation (Pickering & Garrod, 2004, 2013), priming usually occurs because participants are trying to align their representations of the situation under discussion. This type of alignment does not necessarily extend to naturalistic comprehension of stories, where no interlocutor is present. Further studies should investigate priming effects using a variety of texts and modalities (e.g., production or dialogue), so as to determine whether this affects the occurrence of priming effects.

While we suggested that repetition priming may have a limited role in naturalistic comprehension, it must be pointed out that various factors may have impacted our ability to detect priming effects in this study, which may instead emerge when these factors are accounted for. One intriguing hypothesis is that embedding a word in a sentence both at first and subsequent presentation may eliminate repetition priming effects. Indeed, previous studies that investigated the effect of embedding on priming used sentences in the priming phase but presented words in isolation in the test phase (see Besken and Mulligan (2010); Hodapp and Rabovsky (2021); Masson and Macleod (2000)but also Rommers and Federmeier (2018b)). On the other hand, in this study both repeated and non-repeated words were always embedded in sentences.

Furthermore, the distance between successive instances of the same word may also have affected our ability to detect priming effect. Indeed, while a few studies found evidence of very long-lasting priming effects - up to 48 weeks after the first encounter with a word (Cave, 1997) - priming tends to decrease at increasing distance between prime and target (e.g., Durso & Johnson, 1979; Mitchell, 1989; Wheeldon & Monsell, 1992). Unlike most studies on repetition priming, which kept the distance between primes and target constant, in our study words could be repeated at any point in the story, which spanned about one hour (Heilbron et al., 2019). It is therefore possible that priming effects did not emerge in the current study due to the distance between repetitions (or to the combined effect of sentential embedding at first and repeated presentation and of distance). In order to determine whether the distance between repetitions may have affected our ability to detect priming effects, we ran a post-hoc exploratory analysis where the EEG signal was modelled against repetition and the interaction between repetition and distance (see Appendix C). However, the results of the TFCE showed that neither the effect of repetition nor that of the interaction were significant, suggesting that factors other than distance may be responsible for the lack of repetition priming in this study.

5.4 Conclusions

In conclusion, while we were able to find some weak evidence of repetition priming, the effect disappeared when we took the effect of predictability into account. We hypothesised that repetition priming effects may be more limited than those of predictability in naturalistic comprehension. However, we also identified a few factors that may have contributed to our failure to detect priming effects - such as the distance between repetitions - and we identified suggestions as to how future work could overcome the limitations of this study.

5.5 Appendix A

Number of words per part of speech	
noun	575
verb	483
adjective	226
adverb	99
name	58
preposition	29
NA's	14
number	11
pronoun	7
determiner	3
conjunction	1
interjection	1
unclassified	1

Table 5.2: The table describes the number of words in the dataset per part of speech. Words classified as NA refer to words that did not have a match in the SUBTLEX-UK dataset and were not included in the analyses.

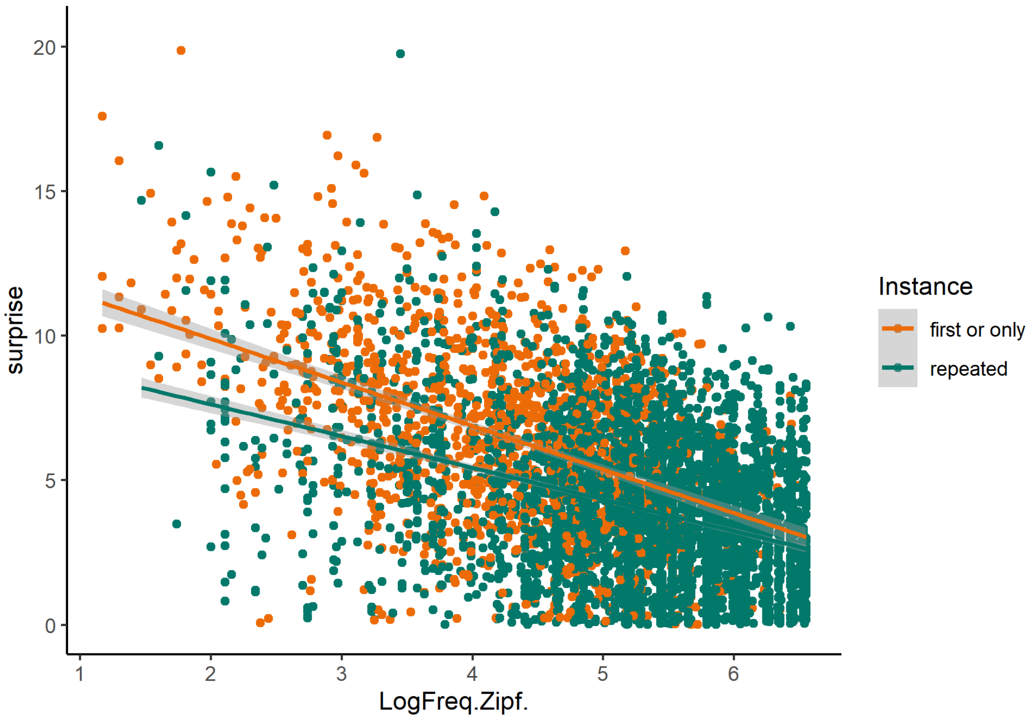

Figure 5.3: Figure plotting frequency values against surprise values for repeated words and unrepeated words/first instances. The regression line was obtained using geom_smooth = “lm” in R.

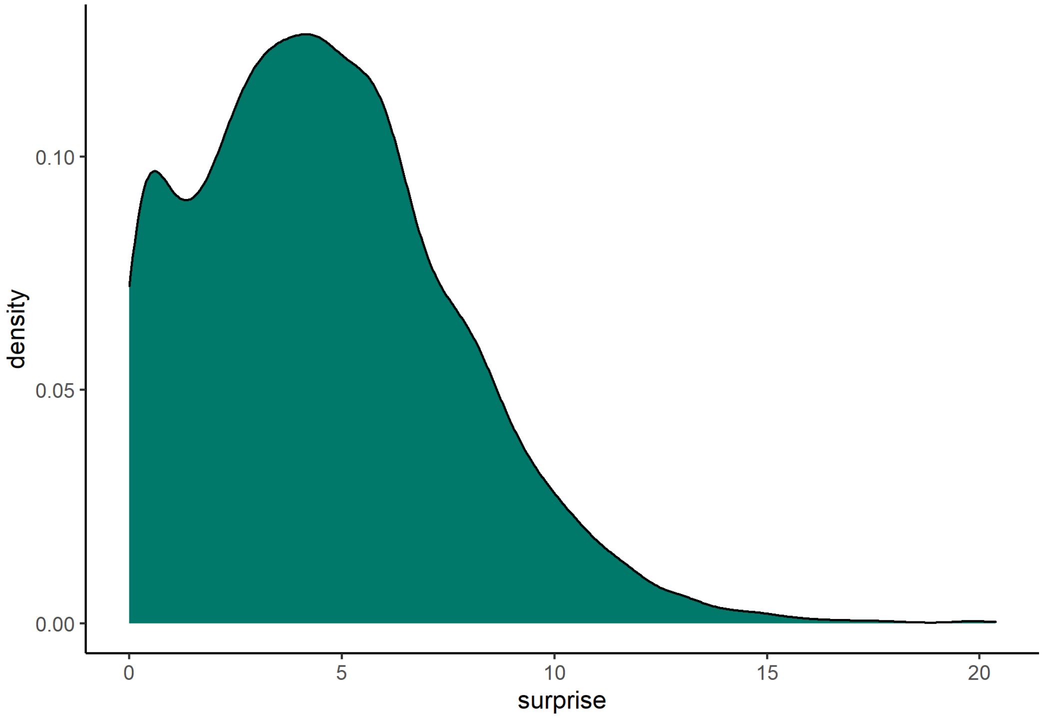

Figure 5.4: Figure plotting the distribution of GPT-2 surprise values for all words in the EEG dataset.

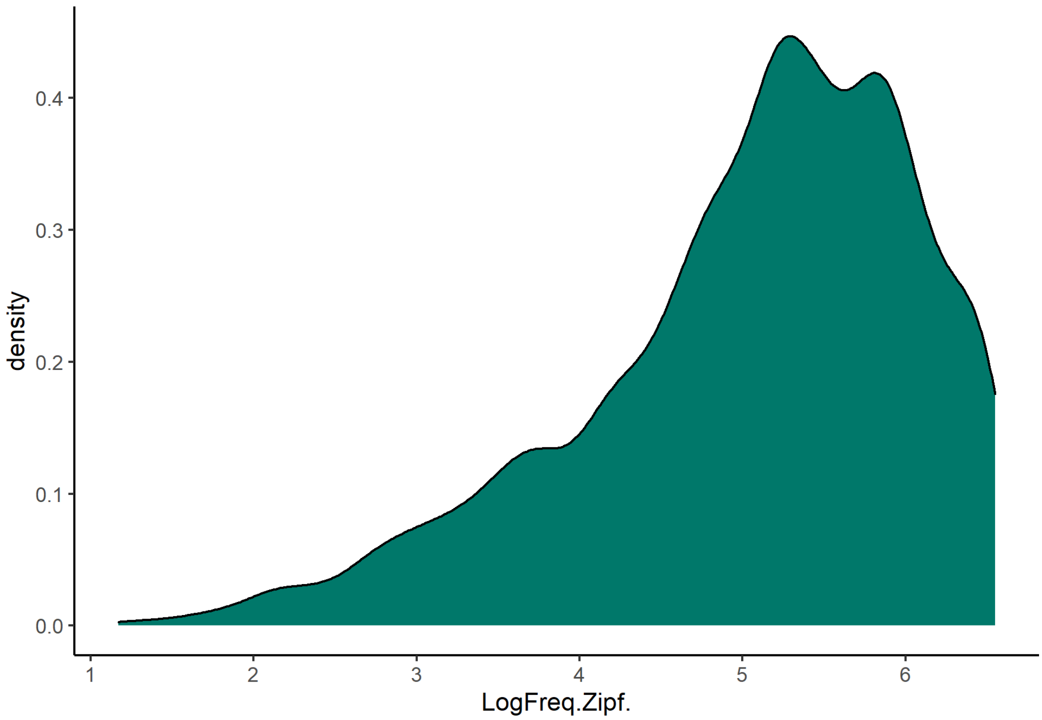

Figure 5.5: Figure plotting the distribution of frequency values for all words in the EEG dataset.

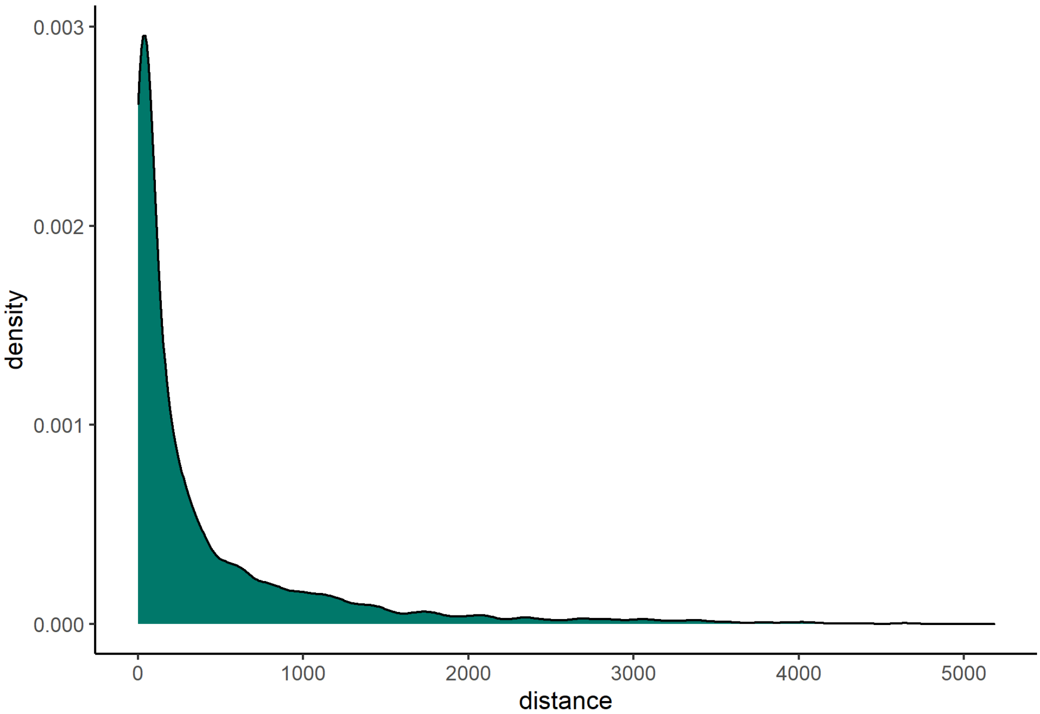

Figure 5.6: Figure plotting the distribution of distance values (i.e., number of intervening words between successive repetitions of a word) in the EEG dataset.

5.6 Appendix B

The figures below show the channel labels and location on the layout used for recording the EEG data analysed in this study and the output of the predictors in the models that were not tested statistically.

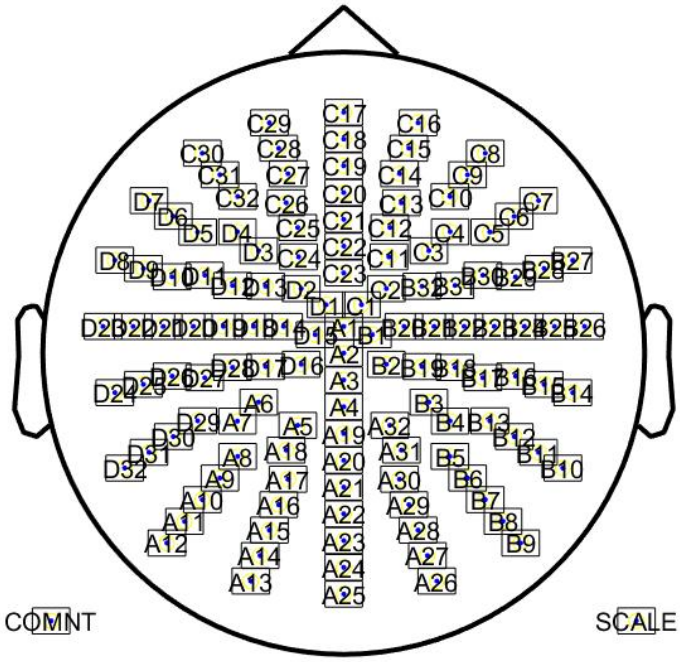

Figure 5.7: Channel labels and locations on the Biosemi 128-electrode layout used for recording the EEG data analysed in this study.

Model 0

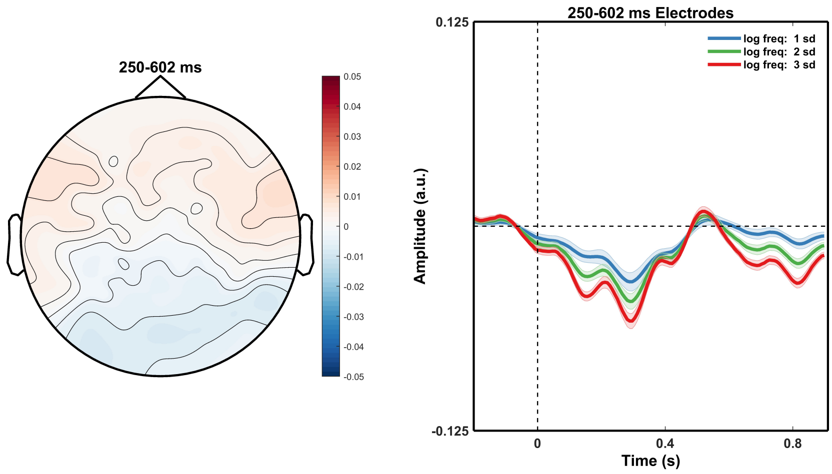

Figure 5.8: Output of frequency in model 0. Left: The topography shows the difference between the averaged coefficients of the predictor for frequency and zero. *Right*: In the model, the intercept captures the effect when the surprise predictor is 0. The waveforms depict the timecourse of frequency at increasing standard deviation from the mean (1SD, 2SD, 3SD). The waveforms were averaged over the same electrodes that contributed to the cluster between 250-602ms in model 1 (A1, A2, A3, A4, A5, A6, A7, A8, A9, A10, A11, A12, A13, A14, A15, A16, A17, A18, A19, A20, A21, A22, A23, A24, A25, A26, A27, A28, A29, A30, A31, A32, B1, B2, B3, B4, B5, B6, B7, B8, B9, B10, B11, B12, B13, B16, B17, B18, B19, D14, D15, D16, D25, D26, D27, D28).

Model 1

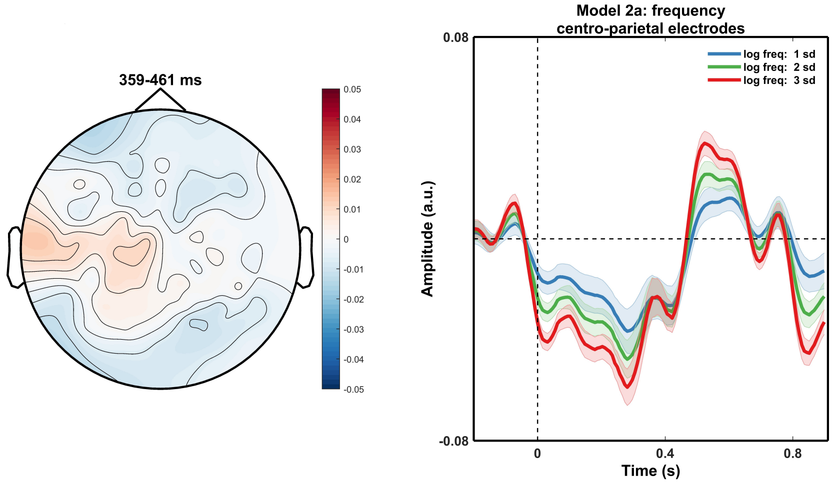

Figure 5.9: Output of frequency in model 1. Left: The topography shows the difference between the averaged coefficients of the predictor for frequency and zero. *Right*: In the model, the intercept captures the effect when the predictor for frequency is 0. The waveforms depict the timecourse of the sum of the intercept and the marginal coefficients for frequency at increasing standard deviation from the mean (1SD, 2SD, 3SD). The waveforms were averaged over the same electrodes that contributed to the cluster between 250-602ms in model 1 (A1, A2, A3, A4, A5, A6, A7, A8, A9, A10, A11, A12, A13, A14, A15, A16, A17, A18, A19, A20, A21, A22, A23, A24, A25, A26, A27, A28, A29, A30, A31, A32, B1, B2, B3, B4, B5, B6, B7, B8, B9, B10, B11, B12, B13, B16, B17, B18, B19, D14, D15, D16, D25, D26, D27, D28).

Model 2

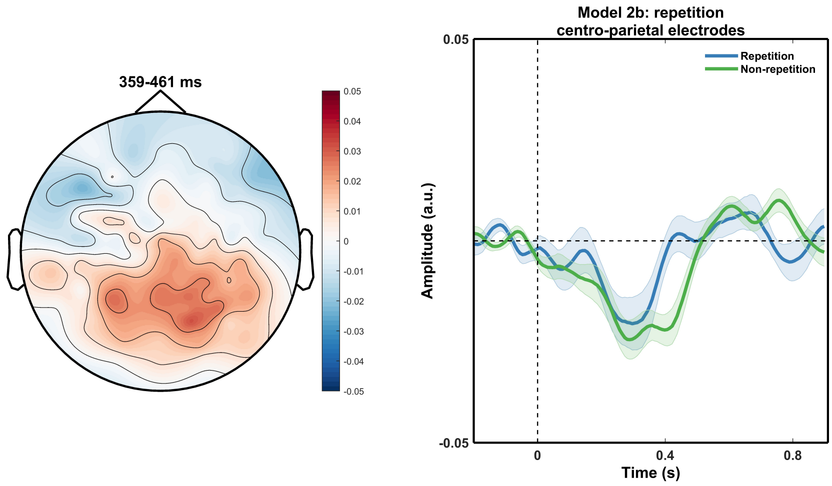

Figure 5.10: Output of repetition in model 2. Left: The topography shows the difference between the averaged coefficients of the predictor for repetition and zero. *Right*: The waveforms depict the timecourse of the intercept (non-repetition) and the sum of the intercept and the coefficients of repetition. The waveforms were averaged over the same electrodes that contributed to the cluster between 250-602ms in model 1 (A1, A2, A3, A4, A5, A6, A7, A8, A9, A10, A11, A12, A13, A14, A15, A16, A17, A18, A19, A20, A21, A22, A23, A24, A25, A26, A27, A28, A29, A30, A31, A32, B1, B2, B3, B4, B5, B6, B7, B8, B9, B10, B11, B12, B13, B16, B17, B18, B19, D14, D15, D16, D25, D26, D27, D28).

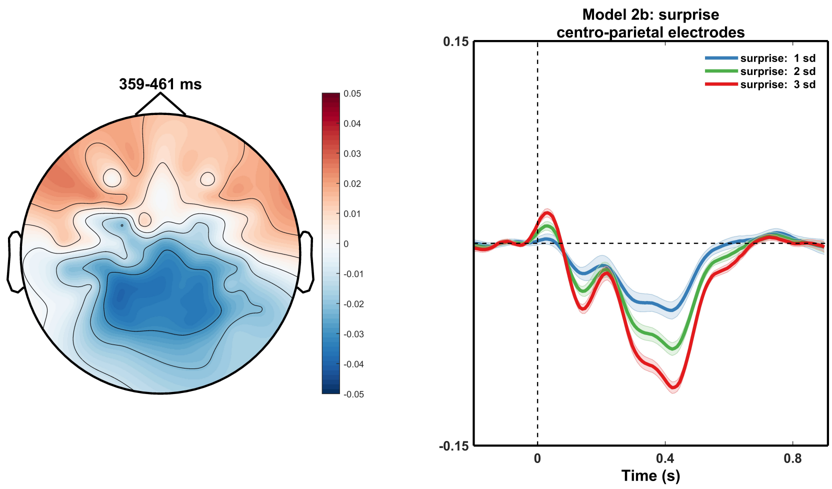

Figure 5.11: Output of surprise in model 2. Left: The topography shows the difference between the averaged coefficients of the predictor for surprise and zero. *Right*: The waveforms depict the timecourse of the sum of the intercept and the marginal coefficients for surprise at increasing standard deviation from the mean (1SD, 2SD, 3SD). The waveforms were averaged over the same electrodes that contributed to the cluster between 250-602ms in model 1 (A1, A2, A3, A4, A5, A6, A7, A8, A9, A10, A11, A12, A13, A14, A15, A16, A17, A18, A19, A20, A21, A22, A23, A24, A25, A26, A27, A28, A29, A30, A31, A32, B1, B2, B3, B4, B5, B6, B7, B8, B9, B10, B11, B12, B13, B16, B17, B18, B19, D14, D15, D16, D25, D26, D27, D28).

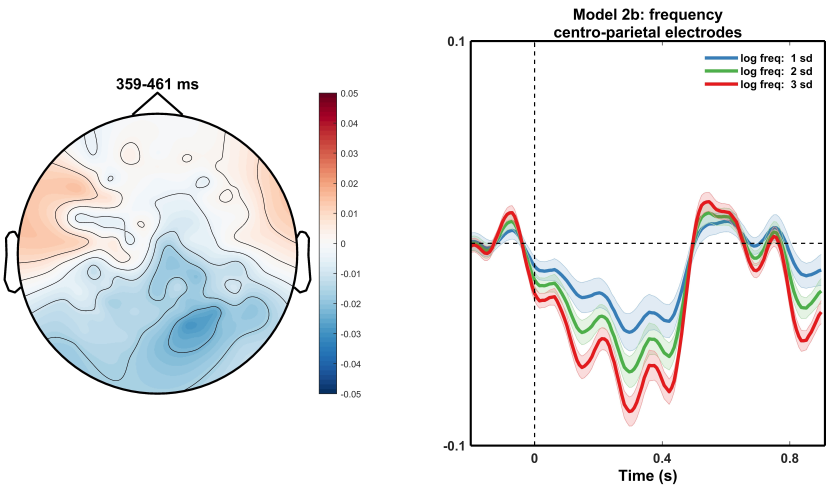

Figure 5.12: Output of frequency in model 2. Left: The topography shows the difference between the averaged coefficients of the predictor for frequency and zero. *Right*: The waveforms depict the timecourse of the sum of the intercept and the marginal coefficients for frequency at increasing standard deviation from the mean (1SD, 2SD, 3SD). The waveforms were averaged over the same electrodes that contributed to the cluster between 250-602ms in model 1 (A1, A2, A3, A4, A5, A6, A7, A8, A9, A10, A11, A12, A13, A14, A15, A16, A17, A18, A19, A20, A21, A22, A23, A24, A25, A26, A27, A28, A29, A30, A31, A32, B1, B2, B3, B4, B5, B6, B7, B8, B9, B10, B11, B12, B13, B16, B17, B18, B19, D14, D15, D16, D25, D26, D27, D28).

5.7 Appendix C

In the main analyses reported for this study, we compared the first or only instance of each word (unrepeated words) against all repetitions. While this approach provides a good approximation to what has been done in previous repetition priming studies, one downside is that a different number of repeated and unrepeated words were present in the models. This is typically well matched in lab experiments. We therefore ran additional models to determine whether the absence of priming effects in our experiment may have been due to this imbalance. In these models, we included only the first (unrepeated) and second (repeated) occurrence of each word. The dataset included 697 first occurrences and 700 second occurrences (706 first and second occurrences before removing the first and last word token in each run). The models tested whether the EEG signal was predicted by repetition, word frequency, the interaction between word frequency and repetition, and surprise (models 1A, 2A, and 3A). Unlike in the models from the main analyses, we additionally investigated whether the distance between the first and second occurrence of a word was predictive of the N400 amplitude (model 4A). In model 4A, we therefore included the interaction between distance and repetition. Distance was calculated as the number of intervening words between the first and second occurrence: words that were repeated immediately after the first occurrence were assigned distance = 1, while first occurrences were always assigned distance = 0.

As a last post-hoc exploratory analysis, we ran two models using the same predictors as model 2A in the main analysis but, unlike in the main analyses, we calculated repetitions within runs (models 5A and 6A below). While first instances and repetitions across runs may be separated by a wide range of intervening items, calculating repetitions within runs allows to minimise the effect of distance. In model 5A, we compared the first or only instance of each word (unrepeated words) against all repetitions. In model 6A, we only compared first and second occurrences. The dataset used in model 6 included 932 first occurrences and 932 second occurrences (after removing the first and last word token in each run).

Below we report the details of each model and the results of the TFCE analyses:

Model 1A

$$EEG \sim 1, EEG \sim cat(repetition)$$

In the TFCE analysis, we tested the effect of the predictor for repetition, which was not statistically significant ($p > .05$).

Model 2A

$$EEG \sim 1, EEG \sim cat(repetition) * frequency$$

In the TFCE analysis, we tested the effect of the predictors for repetition, frequency and their interaction. None of the predictors were statistically significant ($p > .05$).

Model 3A

$$EEG \sim 1, EEG \sim cat(repetition) * frequency + surprise$$

In the TFCE analysis, we tested the effect of the predictors for repetition, frequency, their interaction, and surprise. The predictors for surprise and frequency were statistically significant ($p < .05$), while the other predictors were not ($p > .05$). The significant clusters in the TFCE for surprise and frequency are shown in the table below. Concerning the effect of frequency, the effect was elicited by only two channels and lasted for < 50ms. For this reason, we do not consider the effect meaningful. The surprise effect recapitulates the relationship between predictability and the N400 identified for model 0 in the analyses from the main text.

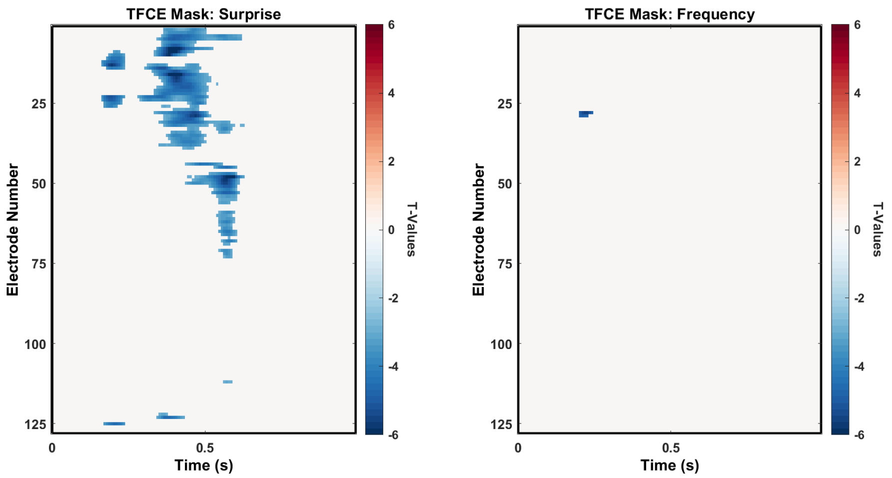

Figure 5.13: Effect of frequency and surprise in model 3A. The figures show significant clusters in the TFCE analysis for the predictors surprise (left) and word frequency (right).

Model 4A

$$EEG \sim 1, EEG \sim cat(repetition) + cat(repetition) : distance$$

In the TFCE analysis, we tested the effect of the predictors for repetition and the interaction between repetition and distance. None of the predictors were statistically significant ($p > .05$). We conclude that the instability of repetition effects observed in the main text cannot be attributed to an imbalance between the number of repeated and unrepeated words in the models, as we do not observe any relationship between repetition and the N400 in the analyses reported here when equating the number of repeated and unrepeated words.

Models 5A and 6A

$$EEG \sim 1, EEG \sim cat(repetition) * frequency + surprise$$

In the TFCE analysis, we tested the effect of the predictors for repetition, surprise, frequency, and the interaction between repetition and frequency. In model 5A (including first or only instances of a word vs all repetitions), all predictors were significant ($p > .05$). In model 6A (inluding only first and second instances of a word), the effects of the predictors for surprise and frequency were statistically significant ($p < .05$), but the effects of the predictors for repetition and the interaction between frequency and repetition were not ($p > .05$).

The significant clusters in the TFCE for each predictor are shown in Figure 5.14 (model 5A) and 5.16 (model 6A). Concerning the effect of repetition in model 5A, the TFCE yielded a negative cluster between around 200 and 400ms with a frontal distribution (see Figure 5.15), suggesting that repeated words were associated with a more negative waveform than unrepeated words. The direction and topography of the effect are not consistent with a classic N400, which has a centro-parietal distribution and is usually more negative for non-repeated words than repeated words. In model 5A, the topography and timing of the effect are more consistent with a N200, which has been associated with cognitive control processes (Folstein & Van Petten, 2008) and has also been found in studies on negative priming, i.e., slower or less accurate processing of repeated stimuli that have been previously ignored (Frings & Groh-Bordin, 2007). While the functional meaning of the N200 effect in this study is unclear, the effect was not present in model 6A. For this reason, we are cautious in interpreting the effect. Regardless of the N200-like cluster found in model 5A, the results of both model 5A and model 6A showed that repetition is not associated with modulations of the N400. This means that the instability of repetition priming effects in the main analyses did not depend on any discrepancies between the way in which repetition and surprise values were calculated (i.e., within runs or across runs).

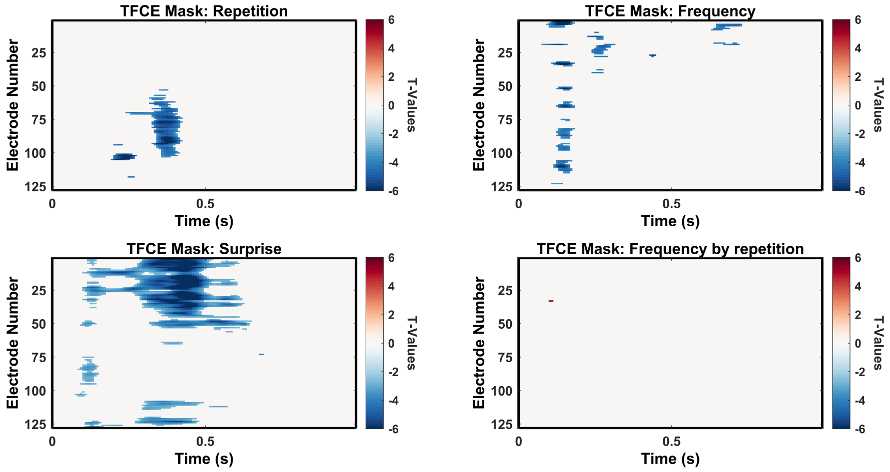

Figure 5.14: Effect of repetition, frequency, surprise, and the interaction between frequency and repetition in model 5A. The figures show significant clusters in the TFCE analysis for the predictors repetition (top left), word frequency (top right), surprise (bottom left), and the interaction between frequency and repetition (bottom right).

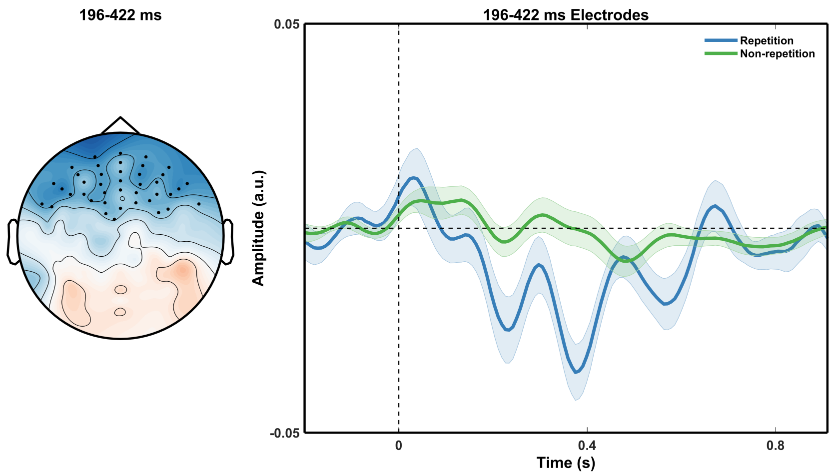

Figure 5.15: Effect of repetition in model 5A. Left: Scalp topographies showing the difference between the averaged coefficients of the predictor for repetition and zero. The black marks indicate electrodes contributing to each cluster (significant at alpha < 0.05, two-tailed). *Right*: In the model, the intercept captures the effect when there is no effect of repetition. The waveforms depict the timecourse of the intercept (non-repetition) and the sum of the intercept and the coefficients of repetition. The waveforms were averaged over the electrodes contributing to the cluster (standard errors of the mean indicated by shaded area around the lineplots).

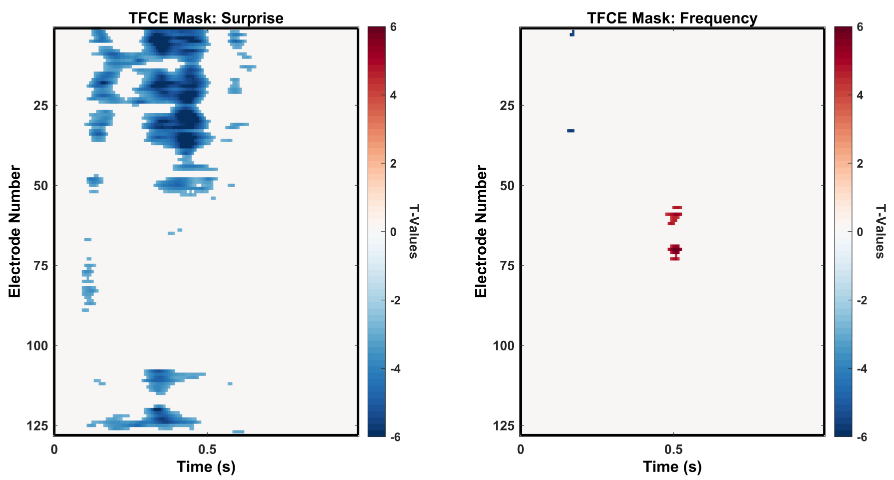

Figure 5.16: Effect of surprise and frequency in model 6A. The figures show significant clusters in the TFCE analysis for the predictors surprise (left), and word frequency (right).

6 | General discussion

6.1 Introduction

In this dissertation, I aimed to explore how language processes unfold during conversation. As one of the research questions, I investigated whether there is interference between comprehension and production processes when they are performed in parallel. This question was motivated by the fact that comprehension and production may partly overlap in conversation, as suggested by the observation that gaps between turns are usually longer than average picture-naming latencies (Levinson, 2016). In the preceding chapters I mainly focused on the interference from production to comprehension and, more specifically, on whether speech planning affects simultaneous semantic processing during comprehension.

In addition to investigating interference between production and comprehension processes, I also asked whether any costs associated with language production and comprehension can be mitigated by priming, which has a pivotal role as supporting mechanism in models of dialogue (e.g., Pickering & Garrod, 2004). In order to determine whether priming can indeed aid language processing in naturalistic settings, I tested the occurrence of this effect under the conditions that prevail in everyday comprehension and/or production. In particular, I investigated whether repetition priming occurs during naturalistic comprehension and during linguistic dual-tasking.

In the following sections, I will briefly summarise the results and findings of the previous chapters and discuss the main findings.

6.2 Summary of results

In *Chapter 2*, I reviewed previous studies on repetition priming in comprehension and production. In particular, I focused on three main factors that might affect repetition priming, namely the distance between prime and target, the embedding of the prime, or the division of attention between production and comprehension. The topics discussed in *Chapter 2* were further explored and investigated in the subsequent chapters. In particular, the influence of lag on repetition priming was further discussed in *Chapter 3*. The question of how and whether embeddings can affect the occurrence and size of the priming effect was investigated extensively in *Chapter 5* and, more briefly, in *Chapter 3*. Finally, *Chapter 3* and *Chapter 4* were entirely dedicated to understanding whether linguistic dual-tasking hinders repetition priming from comprehension to production.

More specifically, in *Chapter 3*, I investigated whether repetition priming is hindered if participants carry out a concurrent production task while listening to the prime word. In two experiments, participants heard sentences containing prime words, and then named two pictures, a distractor and a target. The target could be unrelated, identical to the immediately preceding prime, or to the prime presented ten or fifty trials earlier. In the first experiment, the distractor picture was presented at the end of the sentence, meaning that comprehension of the sentence and planning of the picture name were carried out sequentially. In the second experiment, the distractor picture was presented at the onset of the prime word in the sentence. This means that in Experiment 2 comprehension and speech planning overlapped. Since in Experiment 2 attentional resources had to be split between comprehension and production, I hypothesised that repetition priming effects might be smaller in Experiment 2 than in Experiment 1, or that the priming effect would decay faster across lags in Experiment 2 than in Experiment 1. In Experiment 3 I showed that priming was evident even when only lags of 10 or 50 trials were used, suggesting that the effect did not depend on participants strategically processing the prime to ease task demands.

While the paradigm used in *Chapter 3* did not afford to test whether interference occurred during initial processing of the prime, in *Chapter 4* I used EEG to investigate not only repetition priming but also online comprehension processes during both single-tasking and linguistic dual-tasking. I hypothesised that in the studies reported in *Chapter 3* the lag between the onset of the distractor and the target was long enough for participants to recover from any dual-tasking interference but that effects would appear when using more sensitive measures, such as ERPs. The paradigm used was similar to that of *Chapter 3*, with two main differences. First of all, primes and targets were now split into different phases, a study phase (in which primes were processed while naming a picture or in silence), and a test phase, where participants named identical or unrelated targets. Furthermore, the planning condition was implemented within-participants, with the no plan block always preceding the plan block. I predicted that the N1 and N400 elicited by the prime word in the study phase would have a smaller amplitude in the plan block than in the no plan block, due to the division of attention between production and comprehension. Similarly, I hypothesised that the repetition priming effect in the test phase, indexed by a reduced N400 for primed than unprimed words, would be greater in the plan than in the no plan block. However, none of these predictions were borne out.

In *Chapter 5*, I focused solely on comprehension: I investigated how repetition priming effects unfold in natural story comprehension and whether they are affected by the embedding of the prime word at first presentation. To do so, I analysed the EEG signal recorded while participants listened to a narrative text. In particular, I predicted that repeated words would yield a reduced N400 with respect to unrepeated words. The results from this study were ambiguous: while a first analysis showed that repeated words were indeed associated with a reduced N400 with respect to unrepeated words, repetition priming effects vanished in a subsequent analysis that controlled for the effect of word predictability.

As discussed in *Chapter 5*, it is possible that priming has a more limited role than prediction in naturalistic comprehension. According to this account, prediction is sufficient to guide processing under most circumstances and priming may only be beneficial when prediction fails. More specifically, while initial processing of low-frequency and unpredictable words may be effortful, priming can ease processing costs at subsequent repetition, and therefore support comprehension. Another possibility is that repetition priming may be especially useful in dialogue - when the interlocutors try to align their interpretations of the situation under discussion - but may have a more limited role during naturalist comprehension of stories. It is important to point out that a number of factors may have affected the possibility to detect priming effects in *Chapter 5*, such as the distance between words or the measure used to quantify predictability. Future studies should explore repetition priming in naturalistic settings using a wider variety of texts. For instance, priming effects may emerge when using shorter texts, where the number of intervening words between repetitions is smaller. Alternatively, dual-EEG studies could help assess whether repetition priming occurs when participants are engaged in a conversation.

6.3 Repetition priming as a supporting mechanism during linguistic dual-tasking

Chapter 3 focused on whether repetition priming can ease demands associated with linguistic dual-tasking by presenting prime words (presented auditorily) in or without overlap with a production task. While a previous experiment (Jongman & Meyer, 2017) had already shown that repetition priming from comprehension to production is not affected by a concurrent production task, the authors only focused on priming of single words that were repeated immediately. Yet, in real-life conversations, turns vary in length, from single words to multi-sentence narratives (e.g., Corps, Knudsen, & Meyer, 2022), and therefore it is likely that repeated words are embedded in utterances contexts and that they are not necessarily repeated immediately. The question was therefore whether linguistic dual-tasking would hinder repetition priming when the comprehension task was made more complex. While in *Chapter 3* I embedded primes in sentences and targets were repeated after three different lags, repetition priming was never affected by a concurrent production task.

According to Pickering and Garrod (2004), conversations are easy because interlocutors progressively reach a common representation of the event that is being discussed. This process, called alignment, partly comes about through priming, which occurs at different levels (e.g., phonological, lexical, syntactic level). As described in the *Introduction*, conversation is a complex activity, in which participants continuously switch between comprehension and production, and occasionally perform the two processes in parallel. The results obtained in *Chapter 3* suggest that repetition priming can in principle work as a supporting mechanism in conversation and that any overlap between production and comprehension does not have adverse effects on priming, regardless of whether or not the prime is embedded in a sentence and regardless of the delay. The finding that repetition priming at different lags was not affected by the production task is especially interesting, since repetition priming has been found to decay at different rates depending on various factors, such as the stimulus type and modality of presentation (e.g., words vs nonwords in visual and auditory modalities, McKone & Dennis, 2000). Yet, *Chapter 3* showed that linguistic dual-tasking is not one of these factors, and that the encoding necessary to yield priming effects is not modulated by a secondary linguistic task.

While *Chapter 3* showed that linguistic dual-tasking does not hinder repetition priming, even when the comprehension task is made more complex with respect to previous studies (e.g., Jongman & Meyer, 2017), the paradigm employed is still very distant from real-life conversation. One of the most evident limitations is that, while I tried to increase the complexity of the comprehension and priming tasks by using embeddings and different lags between primes and targets, the production task always required naming a single picture. Previous studies have suggested that any effects of a secondary task on implicit memory tasks (such as fragment completion) might depend on the difficulty of the secondary task (Mulligan, 1997; Wolters & Prinsen, 1997). Thus, in order to determine whether repetition priming is truly resilient to division of attention, future studies should manipulate the complexity of the production task. For instance, participants could be asked to produce more complex descriptions containing adjectives in addition to nouns (e.g., *blue striped shirt*) or to describe events rather than naming simple objects (e.g., *Sue is hugging Hannah*). If repetition priming is truly resilient to linguistic dual-tasking, there should not be any effects of task difficulty on the magnitude of repetition priming effects.

Another limitation of the findings of this chapter concerns the sole focus on repetition priming. Yet, priming in the interactive alignment model does not only refer to repetition of words but also word meanings, syntactic structures, etc. While linguistic dual-tasking does not impact on repetition of words, there may be effects on other types of priming. For instance, Jongman and Meyer (2017) found that linguistic dual-tasking did not have an impact on repetition priming but, in one of the experiments, hindered associative priming. This suggests that, while repetition priming might be resilient to linguistic dual-tasking, that might not be the case for other forms of priming. One important tenet of the interactive alignment model is that alignment percolates across levels, meaning that *alignment at one level enhances alignment at other levels*. For instance, repetition of words enhances syntactic priming, a phenomenon known as lexical boost (Branigan, Pickering, & Cleland, 2000). In addition to investigating whether syntactic priming and other forms of priming are affected by linguistic dual-tasking, further studies should also determine whether any detrimental effects of linguistic dual-tasking on other types of priming are mitigated by phenomena such as the lexical boost.

6.4 Repetition priming as a supporting mechanism during naturalistic comprehension

While in *Chapter 3* I investigated whether repetition priming can ease demands associated with linguistic dual-tasking, in *Chapter 5* I asked whether it can work as a supporting mechanism during naturalistic comprehension of stories. The main prediction of this study was that repeated words would elicit a reduced N400 with respect to unrepeated words. However, the results of the analyses did not confirm this hypothesis. While the effect of repetition priming was unreliable, the amplitude of the N400 was modulated by word predictability.

Unlike previous studies on repetition priming, where the repeated word was always presented in a weakly constraining sentence, in *Chapter 5* both first occurrences and subsequent repetitions had varying predictability values. One possibility is that, upon repetition, predictability of the word within the sentence may override any priming effects. These findings suggest that the local context of a word, e.g., surrounding words and embedding, may be more relevant than across-story repetition.

In the discussion of *Chapter 5*, I argued that prediction may be sufficient to guide comprehension during naturalistic comprehension of stories and that repetition priming may be come relevant under specific circumstances, when prediction fails to aid processing. For instance, this may be the case of unpredictable or very low-frequency words: when such words initially occur in the text, additional resources are allocated for processing, resulting in facilitated processing when the word is repeated. It is important to note that, in the interactive alignment model, priming is a necessary mechanism to ensure that interlocutors align their representations. Alignment may be less relevant during naturalistic comprehension of texts, where no interlocutors are present.

6.5 The effect of concurrent production on online comprehension processes

One of the research questions I addressed in *Chapter 4* is whether linguistic dual-tasking is associated with decrements of comprehension processes. The question was motivated by previous findings that performing a production and a comprehension task simultaneously hindered production processes (e.g., Bögels et al., 2015; Fargier & Laganaro, 2016). While a few studies investigated the converse question, that is to say whether simultaneous production hindered comprehension processes, the results are not conclusive. For this reason, I recorded EEG activity of participants that listened to prime words with or without a concurrent picture-naming task (no plan vs plan block). On the basis of previous studies, I made two main predictions. First of all, the N1 elicited by prime words should be reduced during linguistic dual-tasking with respect to single tasking, due to the fact that attentional resources should be split between production and comprehension. The same pattern should also hold for the N400, which would hint at more shallow processing of comprehended words during linguistic dual-tasking than during single-tasking.

However, none of the predictions were confirmed by the data. First of all, while there was a difference between the N1 in the two blocks, the effect was in the opposite direction with respect to the initial hypothesis: i.e., the amplitude was greater in the plan block than in the no plan block. The cluster-based permutation also highlighted a second cluster in the data. While the topography and timecourse might in principle be consistent with a N400, the morphology of the waveform suggested that the cluster reflected a sustained positivity, which was greater in the plan block than in the no plan block. In the remainder of this section I will first describe some of the hypotheses about the functional meaning of this positivity, and suggest ways to disentangle them. Then, I will describe methodological factors that might affect the ability to detect N1 and N400 differences in studies on the effect of linguistic dual-tasking on comprehension.

In *Chapter 4*, I advanced four main hypotheses about the functional role of the sustained positivity, namely speech planning, increased attention to the cue to speak, or decision processes and response preparation. An explanation in terms of speech planning processes was ruled out because a similar sustained positivity has been found when picture naming is replaced by a button-press task (Jongman et al., 2020). Similarly, I argued that the component should not index working memory processes, since the sustained positivity also occurs in immediate picture-naming tasks, where there is no delay between the onset of the stimulus and the moment participants can utter their response (Eulitz et al., 2000). In the discussion of *Chapter 4*, I therefore restricted the available hypotheses about the functions of the sustained positivity to three: namely, increased attention to the cue to speak, decision or response preparation processes.

One possibility that was not considered in *Chapter 4* is that the sustained positivity reflects attention and processing of the to-be-named stimulus, rather than the cue to speak. Unlike the interpretation of Jongman et al. (2020), namely that the sustained positivity reflects the attention to the cue to speak, this last hypothesis can explain why the sustained positivity occurs in paradigms in which there is no cue to speak, and participants are required to name the picture as soon as it is presented on the screen (e.g., Gerakaki, 2020). Furthermore, in our experiment, the sustained positivity occured both for pictures in the NP block and in the P block (although it was enhanced in the P block). In the NP block, no cue to speak was present and participants were never required to produce a response. For this reason, an explanation in terms of enhanced attention/processing the to-be-named stimulus is more likely with respect to an explanation in terms of enhanced attention to the cue to speak. Furthermore, the former account explains why the topography of the positivity changes when the stimulus is presented in a different modality (Jongman et al., 2020) but not when the response modality changes (e.g., picture naming versus button press).

The paradigm used in *Chapter 4* does not make it possible to distinguish between the hypotheses advanced above (attention to the cue to speak, decision, response preparation processes and attention and processing of the stimulus). One way to determine the functional role of this sustained positivity would be to record participants' EEG activity while they carry out different versions of the picture-naming task. Participants simply look at the picture, without giving an overt response (passive viewing, akin to the NP block in *Chapter 4*), name the picture after a delay, at the onset of the cue to speak (delayed naming, akin to the P block in *Chapter 4*), make a button-press at the onset of the cue (delayed button press), or make a button press at the onset of the picture (immediate button press). If the positivity occurs both in the delayed picture-naming and delayed button-press conditions, one would exclude that the positivity reflects speech planning processes, therefore replicating Jongman et al. (2020). If the positivity also occurs in the passive viewing condition (as in the NP block in *Chapter 4*), the response preparation hypothesis should also be excluded, as no response is required in the passive-viewing condition. Finally, if the positivity occurs in both button press conditions, delayed and immediate, one would argue that the sustained positivity should not index attention to the cue to speak, as the immediate button-press condition (but also the passive viewing condition) does not include a cue to speak. In that case, the most likely explanation for the positivity would be to state that it reflects attention (and, possibly, processing) of the stimulus.

Regardless of the functional role of the sustained positivity, its occurrence in studies of linguistic dual-tasking warrants methodological considerations about the ability to detect any effects of simultaneous speech planning on comprehension processes. Indeed, any influence of speech planning on the amplitude of the N400 might be overridden by the occurrence of the sustained positivity, which has a similar timecourse and topography as the N400. This makes it challenging to use EEG to investigate how semantic processing during comprehension differs between linguistic dual-tasking and single-tasking.

While EEG paradigms may not be suited to exploring how semantic processing during comprehension is affected by a concurrent speech-planning task, they still provide a useful tool to study how early auditory processes are affected by modulations of attention, as indexed by the N1 component. While the results in *Chapter 4* did not confirm the initial predictions about the N1, I argued that the paradigm used might have affected the ability to detect the hypothesised modulations. In order to maximise the possibility to detect any early modulations of speech planning on comprehension processes, future studies should be designed according to a few principles. First of all, items in no plan and plan conditions should be randomised, so as to make sure that any effects do not depend on strategic processing by participants. Furthermore, items in the production and comprehension tasks should be presented at different SOAs, since the distance between them has been shown to affect the occurrence of any modulations of speech planning on comprehension (Fargier & Laganaro, 2016).

6.6 Conclusions

In this dissertation, I investigated whether repetition priming can work as a supporting mechanism in conversation and whether comprehension is hindered by speech planning during linguistic dual-tasking.

From a theoretical point of view, findings from *Chapter 2* and *Chapter 3* suggest that repetition priming is likely to support conversation. Indeed, in the studies described in this chapter, repetition priming persisted across lags - even during linguistic dual-tasking - and can therefore support conversation beyond immediate repetition of words. Furthermore, there is good evidence that priming from comprehension to production is resilient to linguistic dual-tasking. While repetition priming may not be entirely automatic (Prull, 2013; Spataro et al., 2011), the amount of attention allocated to comprehension during speech planning is sufficient to yield priming effects. While repetition priming is likely to support conversation, it is unclear whether other forms of priming are as resilient as repetition priming and whether they can also aid to ease processing demands.

From a methodological point of view, the results of *Chapter 4* and *Chapter 5* provide insights as to whether and how EEG can be used to explore how semantic processing during comprehension unfolds in naturalistic settings (e.g., during linguistic dual-tasking or naturalistic comprehension of a story). A first methodological aspect to take into consideration is that different overlapping processes may be captured by the EEG signal and it may not be always be possible to disentangle them. Furthermore, some ERP components may overshadow other components of interest, therefore making comparisons between conditions difficult. This is the case of the sustained positivity that emerged in the analysis in *Chapter 4*, which overshadowed any N400 differences between conditions. Due to the predominance of ERP components related to other processes (e.g., production tasks), EEG may not be suitable to investigate specific aspects of comprehension (e.g., semantic processing of comprehended speech during linguistic dual-tasking) and other methods such as fMRI, may be preferred (see also Jongman et al. (2020) for similar conclusions in a study investigating speech planning).

A second methodological aspect to consider when using EEG to investigate naturalistic comprehension is that multiple variables may contribute to the modulation of the ERP components of interest. For instance, the lack of repetition priming in *Chapter 5* may be due to the prevailing effect of prediction, which may have overridden any N400 modulations related to word repetition. Alternatively, the predominance of prediction over priming effects, together with the possible deleterious effect of increasing distance between word repetitions, suggests that in naturalistic comprehension local context may be more important than across-story repetition.

References

Abdel Rahman, R., & Melinger, A. (2011). The dynamic microstructure of speech production: semantic interference built on the fly. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(1), 149.

Allum, P. H., & Wheeldon, L. R. (2007). Planning scope in spoken sentence production: the role of grammatical units. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 791.

Almor, A. (2008). Why does language interfere with vision-based tasks? *Experimental Psychology*, *55*(4), 260.

Arnold, J. E. (2010). How speakers refer: The role of accessibility. *Language and Linguistics Compass*, *4*(4), 187–203.

Arnold, J. E. (2016). Explicit and emergent mechanisms of information status. *Topics in cognitive science*, *8*(4), 737–760.

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition*, *76*(1), B13–B26.

Arnold, J. E., & Zerkle, S. A. (2019). Why do people produce pronouns? pragmatic selection vs. rational models. *Language, Cognition and Neuroscience*, *34*(9), 1152–1175.

Audacity-Team. (2012). Audacity team. *Audio editor and recorder*.

Bard, E. G., Aylett, M. P., Trueswell, J., & Tanenhaus, M. (2004). Referential form, word duration, and modeling the listener in spoken dialogue. *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions*, *173*.

Bargh, J. A. (2014). The historical origins of priming as the preparation of behavioral responses: Unconscious carryover and contextual influences of real-world importance. *Social Cognition*, *32*(Supplement), 209–224.

Barry, C., Hirsh, K. W., Johnston, R. A., & Williams, C. L. (2001). Age of acquisition, word frequency, and the locus of repetition priming of picture naming. *Journal of memory and language*, *44*(3), 350–375.

Barry, C., Johnston, R. A., & Wood, R. F. (2006). Effects of age of acquisition, age, and repetition priming on object naming. *Visual cognition*, *13*(7-8), 911–927.

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.

Barthel, M., Sauppe, S., Levinson, S. C., & Meyer, A. S. (2016). The timing of utterance planning in task-oriented dialogue: Evidence from a novel list-completion paradigm. *Frontiers in psychology*, *7*, 1858.

Bartolozzi, F., Jongman, S. R., & Meyer, A. S. (2021). Concurrent speech planning does not eliminate repetition priming from spoken words: Evidence from linguistic dual-tasking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(3), 466.

Bassili, J. N., Smith, M. C., & MacLeod, C. M. (1989). Auditory and visual word-stem completion: Separating data-driven and conceptually driven processes. *The Quarterly Journal of Experimental Psychology*, *41*(3), 439–453.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

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.

Belke, E., & Stielow, A. (2013). Cumulative and non-cumulative semantic interference in object naming: Evidence from blocked and continuous manipulations of semantic context. *Quarterly Journal of Experimental Psychology*, *66*(11), 2135–2160.

Besken, M., & Mulligan, N. W. (2010). Context effects in auditory implicit memory. *Quarterly Journal of Experimental Psychology*, *63*(10), 2012–2030.

Besson, M., Kutas, M., & Petten, C. V. (1992). An event-related potential (erp) analysis of semantic congruity and repetition effects in sentences. *Journal of Cognitive Neuroscience*, *4*(2), 132–149.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive psychology*, *18*(3), 355–387.

Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition*, *104*(3), 437–458.

Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of experimental psychology:*

General, *129*(2), 177.

Boersma, P., & Weenink, D. (2009). *Praat: doing phonetics by computer (version 5.1)[computer program]. retrieved may 1, 2009.*

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.

Bögels, S., Magyari, L., & Levinson, S. C. (2015). Neural signatures of response planning occur midway through an incoming question in conversation. *Scientific reports*, *5*(1), 1–11.

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.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, *75*(2), B13–B25.

Broderick, M. P., Anderson, A. J., Di Liberto, G. M., Crosse, M. J., & Lalor, E. C. (2018). Electrophysiological correlates of semantic dissimilarity reflect the comprehension of natural, narrative speech. *Current Biology*, *28*(5), 803–809.

Broderick, M. P., Anderson, A. J., & Lalor, E. C. (2019). Semantic context enhances the early auditory encoding of natural speech. *Journal of Neuroscience*, *39*(38), 7564–7575.

Brown, A. S., Neblett, D. R., Jones, T. C., & Mitchell, D. B. (1991). Transfer of processing in repetition priming: some inappropriate findings. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(3), 514.

Brown-Schmidt, S. (2012). Beyond common and privileged: Gradient representations of common ground in real-time language use. *Language and Cognitive Processes*, *27*(1), 62–89.

Brown-Schmidt, S., Byron, D. K., & Tanenhaus, M. K. (2005). Beyond salience: Interpretation of personal and demonstrative pronouns. *Journal of Memory and Language*, *53*(2), 292–313.

Brown-Schmidt, S., & Duff, M. C. (2016). *Memory and common ground processes in language use* (Vol. 8) (No. 4). Wiley Online Library.

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science*, *27*(1), 45–50.

Cabeza, R. (1994). A dissociation between two implicit conceptual tests supports the distinction between types of conceptual processing. *Psychonomic*

Bulletin & Review, *1*(4), 505–508.

Camblin, C. C., Ledoux, K., Boudewyn, M., Gordon, P. C., & Swaab, T. Y. (2007). Processing new and repeated names: Effects of coreference on repetition priming with speech and fast rsvp. *Brain Research*, *1146*, 172–184.

Cave, C. B. (1997). Very long-lasting priming in picture naming. *Psychological science*, *8*(4), 322–325.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological review*, *113*(2), 234.

Clark, H. H., & Marshall, C. (1978). Reference diaries. *American Journal of Computational Linguistics*, 59–65.

Clark, H. H., & Marshall, C. R. (1981). Definite knowledge and mutual knowledge.

Cleland, A. A., & Pickering, M. J. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language*, *49*(2), 214–230.

Coane, J. H., & Balota, D. A. (2010). Repetition priming across distinct contexts: Effects of lexical status, word frequency, and retrieval test. *Quarterly Journal of Experimental Psychology*, *63*(12), 2376–2398.

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. *Cognition*, *175*, 77–95.

Corps, R. E., Knudsen, B., & Meyer, A. S. (2022). Overrated gaps: Inter-speaker gaps provide limited information about the timing of turns in conversation. *Cognition*, *223*, 105037.

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and language*, *41*(3), 365–397.

Daliri, A., & Max, L. (2016). Modulation of auditory responses to speech vs. nonspeech stimuli during speech movement planning. *Frontiers in Human Neuroscience*, *10*, 234.

Damian, M. F., Dorjee, D., & Stadthagen-Gonzalez, H. (2011). Long-term repetition priming in spoken and written word production: evidence for a contribution of phonology to handwriting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(4), 813.

Dell, G. S., & Ferreira, V. S. (2016). Thirty years of structural priming: An introduction to the special issue. *Journal of Memory and Language*, *91*, 1–4.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature neuroscience*, *8*(8), 1117–1121.

Di Liberto, G. M., Wong, D., Melnik, G. A., & de Cheveigné, A. (2019). Low-frequency cortical responses to natural speech reflect probabilistic phonotactics. *Neuroimage*, *196*, 237–247.

Dimigen, O., & Ehinger, B. V. (2020). Analyzing combined eye-tracking/eeg experiments with (non) linear deconvolution models. *BioRxiv*, 735530.

Duff, M. C., Hengst, J., Tranel, D., & Cohen, N. J. (2006). Development of shared information in communication despite hippocampal amnesia. *Nature neuroscience*, *9*(1), 140–146.

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). Multipic: A standardized set of 750 drawings with norms for six european languages. *Quarterly Journal of Experimental Psychology*, *71*(4), 808–816.

Durso, F. T., & Johnson, M. K. (1979). Facilitation in naming and categorizing repeated pictures and words. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(5), 449.

Ehinger, B. V., & Dimigen, O. (2019). Unfold: An integrated toolbox for overlap correction, non-linear modeling, and regression-based eeg analysis. *PeerJ*, *7*, e7838.

Eulitz, C., Hauk, O., & Cohen, R. (2000). Electroencephalographic activity over temporal brain areas during phonological encoding in picture naming. *Clinical neurophysiology*, *111*(11), 2088–2097.

Fairs, A., Bögels, S., & Meyer, A. S. (2018). Dual-tasking with simple linguistic tasks: Evidence for serial processing. *Acta psychologica*, *191*, 131–148.

Fargier, R., & Laganaro, M. (2016). Neurophysiological modulations of non-verbal and verbal dual-tasks interference during word planning. *PloS one*, *11*(12), e0168358.

Fargier, R., & Laganaro, M. (2019). Interference in speaking while hearing and vice versa. *Scientific reports*, *9*(1), 1–13.

Fawcett, J. M., Quinlan, C. K., & Taylor, T. L. (2012). Interplay of the production and picture superiority effects: A signal detection analysis. *Memory*, *20*(7), 655–666.

Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current directions in psychological science*, *11*(1), 11–15.

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the n2 component of the erp: a review. *Psychophysiology*, *45*(1), 152–170.

Fong, D. C.-L., & Saunders, M. (2011). Lsmr: An iterative algorithm for sparse least-squares problems. *SIAM Journal on Scientific Computing*, *33*(5), 2950–2971.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of experimental psychology: Learning, Memory, and Cognition*, *10*(4), 680.

Francis, W. S., Camacho, A., & Lara, C. (2014). Words translated in sentence contexts produce repetition priming in visual word comprehension and spoken word production. *Memory & cognition*, *42*(7), 1143–1154.

Francis, W. S., Corral, N. I., Jones, M. L., & Sáenz, S. P. (2008). Decomposition of repetition priming components in picture naming. *Journal of Experimental Psychology: General*, *137*(3), 566.

Francis, W. S., & Sáenz, S. P. (2007). Repetition priming endurance in picture naming and translation: Contributions of component processes. *Memory & Cognition*, *35*(3), 481–493.

Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty p3: an event-related brain potential (erp) sign of the brain's evaluation of novelty. *Neuroscience & Biobehavioral Reviews*, *25*(4), 355–373.

Frings, C., & Groh-Bordin, C. (2007). Electrophysiological correlates of visual identity negative priming. *Brain Research*, *1176*, 82–91.

Fukumura, K., & van Gompel, R. P. (2012). Producing pronouns and definite noun phrases: Do speakers use the addressee's discourse model? *Cognitive Science*, *36*(7), 1289–1311.

Gabrieli, J. D., Vaidya, C. J., Stone, M., Francis, W. S., Thompson-Schill, S. L., Fleischman, D. A., ... Wilson, R. S. (1999). Convergent behavioral and neuropsychological evidence for a distinction between identification and production forms of repetition priming. *Journal of Experimental Psychology: General*, *128*(4), 479.

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, *27*(2), 181–218.

Garrod, S., & Pickering, M. J. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science*, *1*(2), 292–304.

Garrod, S., & Pickering, M. J. (2013a). Interactive alignment and prediction in dialogue. *Alignment in communication: Towards a new theory of communi-*

cation, 193–204.

Garrod, S., & Pickering, M. J. (2013b). Interactive alignment and prediction in dialogue. *Alignment in communication: Towards a new theory of communication*, 193–204.

Gerakaki, S. (2020). *The moment in between: Planning speech while listening* (Unpublished doctoral dissertation). [Sl: sn].

Gibson, J. M., & Bahrey, R. (2005). Modality-specificity effects in priming of visual and auditory word-fragment completion. *The journal of general psychology*, *132*(2), 117–138.

Glaser, W. R., & Düngelhoff, F.-J. (1984). The time course of picture-word interference. *Journal of experimental Psychology: Human perception and performance*, *10*(5), 640.

Hagoort, P. (2005). On broca, brain, and binding: a new framework. *Trends in cognitive sciences*, *9*(9), 416–423.

Hagoort, P. (2013). Muc (memory, unification, control) and beyond. *Frontiers in psychology*, *4*, 416.

Hagoort, P., & Brown, C. M. (2000). Erp effects of listening to speech: semantic erp effects. *Neuropsychologia*, *38*(11), 1518–1530.

Hamburger, M., & Slowiaczek, L. M. (1996). Phonological priming reflects lexical competition. *Psychonomic Bulletin & Review*, *3*(4), 520–525.

Hamburger, M., & Slowiaczek, L. M. (1998). Repetition priming and experimental context effects. *The American journal of psychology*, *111*(1), 1.

Hartsuiker, R. J., Bernolet, S., Schoonbaert, S., Speybroeck, S., & Vanderelst, D. (2008). Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. *Journal of Memory and Language*, *58*(2), 214–238.

Heath, S., McMahon, K., Nickels, L., Angwin, A., MacDonald, A., van Hees, S., ... Copland, D. (2012a). The neural correlates of picture naming facilitated by auditory repetition. *BMC neuroscience*, *13*(1), 1–13.

Heath, S., McMahon, K., Nickels, L., Angwin, A., MacDonald, A., van Hees, S., ... Copland, D. (2012b). Priming picture naming with a semantic task: an fmri investigation. *PLoS One*, *7*(3), e32809.

Heilbron, M., Armeni, K., Schoffelen, J.-M., Hagoort, P., & de Lange, F. P. (2021). A hierarchy of linguistic predictions during natural language comprehension. *bioRxiv*, 2020–12.

Heilbron, M., Ehinger, B., Hagoort, P., & De Lange, F. P. (2019). Tracking naturalistic linguistic predictions with deep neural language models. *arXiv* *preprint arXiv:1909.04400*.

Henson, R., Rylands, A., Ross, E., Vuilleumeir, P., & Rugg, M. D. (2004). The effect of repetition lag on electrophysiological and haemodynamic correlates of visual object priming. *Neuroimage*, *21*(4), 1674–1689.

Hernandez, A. E., & Reyes, I. (2002). Within-and between-language priming differ: evidence from repetition of pictures in spanish-english bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(4), 726.

Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*(4108), 177–180.

Hink, R. F., Van Voorhis, S., Hillyard, S., & Smith, T. (1977). The division of attention and the human auditory evoked potential. *Neuropsychologia*, *15*(4-5), 597–605.

Hodapp, A., & Rabovsky, M. (2021). The n400 erp component reflects an error-based implicit learning signal during language comprehension. *European Journal of Neuroscience*, *54*(9), 7125–7140.

Horton, W. S. (2005). Conversational common ground and memory processes in language production. *Discourse Processes*, *40*(1), 1–35.

Horton, W. S. (2007). The influence of partner-specific memory associations on language production: Evidence from picture naming. *Language and cognitive processes*, *22*(7), 1114–1139.

Horton, W. S., & Gerrig, R. J. (2016). Revisiting the memory-based processing approach to common ground. *Topics in Cognitive Science*, *8*(4), 780–795.

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition*, *100*(3), 464–482.

Hughes, J. W., & Schnur, T. T. (2017). Facilitation and interference in naming: A consequence of the same learning process? *Cognition*, *165*, 61–72.

Iii, H. L. R., Gallo, D. A., & Geraci, L. (2002). Processing approaches to cognition: The impetus from the levels-of-processing framework. *Memory*, *10*(5-6), 319–332.

Indefrey, P. (2011). The spatial and temporal signatures of word production components: a critical update. *Frontiers in psychology*, *2*, 255.

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1-2), 101–144.

Jacobs, C. L., Yiu, L. K., Watson, D. G., & Dell, G. S. (2015). Why are repeated words produced with reduced durations? evidence from inner speech and homophone production. *Journal of Memory and Language, 84*, 37–48.

Jacoby, L. L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of verbal learning and verbal behavior*, *22*(5), 485–508.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*(3), 306.

Johnson, C. J., Paivio, A., & Clark, J. M. (1996). Cognitive components of picture naming. *Psychological bulletin*, *120*(1), 113.

Jongman, S. R., & Meyer, A. S. (2017). To plan or not to plan: Does planning for production remove facilitation from associative priming? *Acta psychologica*, *181*, 40–50.

Jongman, S. R., Piai, V., & Meyer, A. S. (2020). Planning for language production: The electrophysiological signature of attention to the cue to speak. *Language, Cognition and Neuroscience, 35*(7), 915–932.

Joos, K., Gilles, A., Van de Heyning, P., De Ridder, D., & Vanneste, S. (2014). From sensation to percept: the neural signature of auditory event-related potentials. *Neuroscience & Biobehavioral Reviews*, *42,* 148–156.

Kahn, J. M., & Arnold, J. E. (2015). Articulatory and lexical repetition effects on durational reduction: Speaker experience vs. common ground. *Language, Cognition and Neuroscience*, *30*(1-2), 103–119.

Keane, M. M., Cruz, M. E., & Verfaellie, M. (2015). Attention and implicit memory: priming-induced benefits and costs have distinct attentional requirements. *Memory & cognition*, *43*(2), 216–225.

Kennison, S. M., & Gordon, P. C. (1997). Comprehending referential expressions during reading: Evidence from eye tracking. *Discourse Processes*, *24*(2-3), 229–252.

Keuleers, E., Brysbaert, M., & New, B. (2010). Subtlex-nl: A new measure for dutch word frequency based on film subtitles. *Behavior research methods*, *42*(3), 643–650.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in cognitive sciences*, *4*(12), 463–470.

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.

Lai, M. K., Rommers, J., & Federmeier, K. D. (2021). The fate of the unexpected: Consequences of misprediction assessed using erp repetition effects. *Brain*

Research, *1757*, 147290.

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics:(de) constructing the n400. *Nature reviews neuroscience*, *9*(12), 920–933.

Levelt, W. J. (1993). *Speaking: From intention to articulation*. MIT press.

Levinson, S. C. (2016). Turn-taking in human communication–origins and implications for language processing. *Trends in cognitive sciences*, *20*(1), 6–14.

Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in psychology*, *6*, 731.

Levy, B. A., & Kirsner, K. (1989). Reprocessing text: Indirect measures of word and message level processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(3), 407.

Lindsay, L., Gambi, C., & Rabagliati, H. (2019). Preschoolers optimize the timing of their conversational turns through flexible coordination of language comprehension and production. *Psychological science*, *30*(4), 504–515.

Luck, S. J., & Kappenman, E. S. (2011). *The oxford handbook of event-related potential components*. Oxford university press.

Luo, Y.-j., & Wei, J.-h. (1999). Cross-modal selective attention to visual and auditory stimuli modulates endogenous erp components. *Brain research*, *842*(1), 30–38.

MacDonald, P. A., & MacLeod, C. M. (1998). The influence of attention at encoding on direct and indirect remembering. *Acta Psychologica*, *98*(2-3), 291–310.

MacLeod, C. M. (1989). Word context during initial exposure influences degree of priming in word fragment completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(3), 398.

MacLeod, C. M., & Bodner, G. E. (2017). The production effect in memory. *Current Directions in Psychological Science*, *26*(4), 390–395.

MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(3), 671.

Magyari, L., De Ruiter, J. P., & Levinson, S. C. (2017). Temporal preparation for speaking in question-answer sequences. *Frontiers in Psychology*, *8*, 211.

Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language*, *91*, 5–27.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of eeg-and

meg-data. *Journal of neuroscience methods*, *164*(1), 177–190.
Martin, C. D., Branzi, F. M., & Bar, M. (2018). Prediction is production: The missing link between language production and comprehension. *Scientific reports*, *8*(1), 1–9.
Masson, M. E., & Macleod, C. M. (2000). Taking the "text" out of context effects in repetition priming of word identification. *Memory & Cognition*, *28*(7), 1090–1097.
McKone, E. (1995). Short-term implicit memory for words and nonwords. *Journal of experimental psychology: Learning, memory, and cognition*, *21*(5), 1108.
McKone, E. (1998). The decay of short-term implicit memory: Unpacking lag. *Memory & cognition*, *26*(6), 1173–1186.
McKone, E., & Dennis, C. (2000). Short-term implicit memory: Visual, auditory, and cross-modality priming. *Psychonomic bulletin & review*, *7*(2), 341–346.
McQueen, J. M., & Meyer, A. S. (2019). Key issues and future directions: Towards a comprehensive cognitive architecture for language use. In *Human language: From genes and brain to behavior* (pp. 85–96). MIT Press.
Mitchell, D. B. (1989). How many memory systems? evidence from aging. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(1), 31.
Mitchell, D. B. (2006). Nonconscious priming after 17 years: invulnerable implicit memory? *Psychological Science*, *17*(11), 925–929.
Mitchell, D. B., & Brown, A. S. (1988). Persistent repetition priming in picture naming and its dissociation from recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(2), 213.
Monsell, S., Matthews, G., & Miller, D. (1992). Repetition of lexicalization across languages: A further test of the locus of priming. *The Quarterly Journal of Experimental Psychology Section A*, *44*(4), 763–783.
Mulligan, N. W. (1997). Attention and implicit memory tests: The effects of varying attentional load on conceptual priming. *Memory & Cognition*, *25*(1), 11–17.
Mulligan, N. W. (1998). The role of attention during encoding in implicit and explicit memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(1), 27.
Mulligan, N. W., Duke, M., & Cooper, A. W. (2007). The effects of divided attention on auditory priming. *Memory & Cognition*, *35*(6), 1245–1254.
Mulligan, N. W., & Stone, M. (1999). Attention and conceptual priming: Limits

on the effects of divided attention in the category–exemplar production task. *Journal of Memory and Language*, *41*(2), 253–280.

Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic erp analyses: a step-by-step tutorial review. *Brain topography*, *20*(4), 249–264.

Neisser, U. (1954). An experimental distinction between perceptual process and verbal response. *Journal of Experimental Psychology: General*, *47*(6).

Nicolas, S., & Söderlund, H. (2000). Cross-modality priming for individual words in memory for coherent texts. *Scandinavian journal of psychology*, *41*(2), 123–132.

Nieuwland, M. S. (2019). Do 'early'brain responses reveal word form prediction during language comprehension? a critical review. *Neuroscience & Biobehavioral Reviews*, *96*, 367–400.

Oliphant, G. W. (1983). Repetition and recency effects in word recognition. *Australian Journal of Psychology*, *35*(3), 393–403.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). Fieldtrip: open source software for advanced analysis of meg, eeg, and invasive electrophysiological data. *Computational intelligence and neuroscience*, *2011*.

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, *114*(2), 227–252.

Ostergaard, A. L. (1998). The effects on priming of word frequency, number of repetitions, and delay depend on the magnitude of priming. *Memory & cognition*, *26*(1), 40–60.

Parasuraman, R. (1978). Auditory evoked potentials and divided attention. *Psychophysiology*, *15*(5), 460–465.

Parasuraman, R. (1985). Event-related brain potentials and intermodal divided attention. In *Proceedings of the human factors society annual meeting* (Vol. 29, pp. 971–975).

Parmentier, F. B. (2008). Towards a cognitive model of distraction by auditory novelty: The role of involuntary attention capture and semantic processing. *Cognition*, *109*(3), 345–362.

Pashler, H. (2000). 12 task switching and multitask performance. *Control of cognitive processes*, 277.

Pickering, M. J., & Branigan, H. P. (1999). Syntactic priming in language production. *Trends in cognitive sciences*, *3*(4), 136–141.

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological bulletin*, *144*(10), 1002.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and brain sciences*, *27*(2), 169–190.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and brain sciences*, *36*(4), 329–347.

Polich, J. (2003). Theoretical overview of p3a and p3b. *Detection of change*, 83–98.

Polich, J. (2007). Updating p300: an integrative theory of p3a and p3b. *Clinical neurophysiology*, *118*(10), 2128–2148.

Powell, M. J. (2009). The bobyqa algorithm for bound constrained optimization without derivatives. *Cambridge NA Report NA2009/06, University of Cambridge, Cambridge*, *26*.

Prull, M. W. (2013). Attention and repetition priming in the verb generation task. *Acta psychologica*, *143*(2), 218–226.

Rabovsky, M., Hansen, S. S., & McClelland, J. L. (2018). Modelling the n400 brain potential as change in a probabilistic representation of meaning. *Nature Human Behaviour*, *2*(9), 693–705.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I., et al. (2019). Language models are unsupervised multitask learners. *OpenAI blog*, *1*(8), 9.

Roelofs, A. (2003). Goal-referenced selection of verbal action: modeling attentional control in the stroop task. *Psychological review*, *110*(1), 88.

Rommers, J., & Federmeier, K. D. (2018a). Lingering expectations: A pseudo-repetition effect for words previously expected but not presented. *NeuroImage*, *183*, 263–272.

Rommers, J., & Federmeier, K. D. (2018b). Predictability's aftermath: Downstream consequences of word predictability as revealed by repetition effects. *Cortex*, *101*, 16–30.

Rugg, M. D. (1985). The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology*, *22*(6), 642–647.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of experimental psychology: learning, memory, and cognition*, *13*(3), 501.

Schriefers, H., Meyer, A. S., & Levelt, W. J. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of memory and language*, *29*(1), 86–102.

Schwent, V. L., Hillyard, S. A., & Galambos, R. (1976). Selective attention and the auditory vertex potential. ii. effects of signal intensity and masking noise. *Electroencephalography and clinical neurophysiology*, *40*(6), 615–

622.

Segaert, K., Kempen, G., Petersson, K. M., & Hagoort, P. (2013). Syntactic priming and the lexical boost effect during sentence production and sentence comprehension: An fmri study. *Brain and language*, *124*(2), 174–183.

Segaert, K., Wheeldon, L., & Hagoort, P. (2016). Unifying structural priming effects on syntactic choices and timing of sentence generation. *Journal of Memory and Language*, *91*, 59–80.

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.

Smith, M., & Wheeldon, L. (1999). High level processing scope in spoken sentence production. *Cognition*, *73*(3), 205–246.

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage*, *44*(1), 83–98.

Spataro, P., Cestari, V., & Rossi-Arnaud, C. (2011). The relationship between divided attention and implicit memory: A meta-analysis. *Acta psychologica*, *136*(3), 329–339.

Speelman, C. P., Simpson, T. A., & Kirsner, K. (2002). The unbearable lightness of priming. *Acta Psychologica*, *111*(2), 191–204.

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.

Team, R. C. (2020). *R 4.0. 2.* R Foundation for Statistical Computing Vienna, Austria.

Traxler, M. J., Tooley, K. M., & Pickering, M. J. (2014). Syntactic priming during sentence comprehension: evidence for the lexical boost. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 905.

Tsuboi, N., Francis, W. S., & Jameson, J. T. (2021). How word comprehension exposures facilitate later spoken production: implications for lexical processing and repetition priming. *Memory*, *29*(1), 39–58.

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-uk: A new and improved word frequency database for british english. *Quarterly journal of experimental psychology*, *67*(6), 1176–1190.

Van Turennout, M., Bielamowicz, L., & Martin, A. (2003). Modulation of neural activity during object naming: effects of time and practice. *Cerebral Cortex*, *13*(4), 381–391.

van Turennout, M., Ellmore, T., & Martin, A. (2000). Long-lasting cortical plasticity in the object naming system. *Nature neuroscience*, *3*(12), 1329–1334.

Versace, R., & Nevers, B. (2003). Word frequency effect on repetition priming as a function of prime duration and delay between the prime and the target. *British Journal of Psychology*, *94*(3), 389–408.

Vitkovitch, M., Rutter, C., & Read, A. (2001). Inhibitory effects during object name retrieval: The effect of interval between prime and target on picture naming responses. *British journal of psychology*, *92*(3), 483–506.

Walker, M. P., & Stickgold, R. (2004). Sleep-dependent learning and memory consolidation. *Neuron*, *44*(1), 121–133.

Wheeldon, L. R., & Monsell, S. (1992). The locus of repetition priming of spoken word production. *The Quarterly Journal of Experimental Psychology Section A*, *44*(4), 723–761.

Wilder, R. J., Davies, A. G., & Embick, D. (2019). Differences between morphological and repetition priming in auditory lexical decision: Implications for decompositional models. *Cortex*, *116*, 122–142.

Wolters, G., & Prinsen, A. (1997). Full versus divided attention and implicit memory performance. *Memory & Cognition*, *25*(6), 764–771.

Wood, N. L., Stadler, M. A., & Cowan, N. (1997). Is there implicit memory without attention? a reexamination of task demands in eich's (1984) procedure. *Memory & Cognition*, *25*(6), 772–779.

Worringer, B., Langner, R., Koch, I., Eickhoff, S. B., Eickhoff, C. R., & Binkofski, F. C. (2019). Common and distinct neural correlates of dual-tasking and task-switching: a meta-analytic review and a neuro-cognitive processing model of human multitasking. *Brain Structure and Function*, *224*(5), 1845–1869.

Zormpa, E., Brehm, L. E., Hoedemaker, R. S., & Meyer, A. S. (2019). The production effect and the generation effect improve memory in picture naming. *Memory*, *27*(3), 340–352.

Nederlandse samenvatting

Een gesprek voeren is een complexe taak waarbij je meerdere, soms overlappende, processen gebruikt. Wanneer iemand bijvoorbeeld een simpele vraag stelt, moet de gesprekspartner eerst de stroom van geluiden opsplitsen in afzonderlijke woorden en vervolgens aan de slag met hun betekenis en hun relatie tot de andere woorden in de vraag. Om de vraag goed te kunnen beantwoorden, moet de gesprekspartner een specifieke reeks woorden ophalen uit zijn of haar geheugen, deze samenvoegen en vervolgens uitspreken. Het is ook belangrijk dat het antwoord op de vraag op het juiste moment wordt uitgesproken. Wacht je te lang, dan kan dit duiden op aarzeling, maar onderbreek je de spreker terwijl die nog aan het praten is, dan wordt dit vaak als onbeleefd gezien. Hoe kan het dan dat gesprekken zo vanzelf lijken te gaan terwijl er zich op de achtergrond een complexe reeks aan processen afspeelt. Volgens een invloedrijk theoretisch model verlopen gesprekken zo moeiteloos omdat gesprekspartners hun taal op elkaar af gaan stemmen. Dit doen ze door hun eigen uitingen en die van hun gesprekspartner te herhalen (zoals woorden of grammaticale constructies). Dit fenomeen wordt 'priming' genoemd. Met name de priming van woorden, repetitie-priming, is uitvoerig bestudeerd in de psycholinguïstiek. In experimenten waarin mensen worden gevraagd een afbeelding te benoemen of een woord te begrijpen, reageren ze vaak sneller en/of correcter als ze het woord al eerder hebben gehoord of gelezen. Hoewel repetitie-priming een belangrijk en krachtig mechanisme lijkt te zijn, is er nog maar weinig onderzoek gedaan naar de invloed ervan in gesprekssituaties. Vele onderzoeken op het gebied van priming zijn bijvoorbeeld gebaseerd op de directe herhaling van losse woorden. In een gesprek is de situatie vaak anders. Woorden maken deel uit van zinnen en kunnen dus pas na een tijdje worden herhaald. Bovendien moeten deelnemers soms twee taken tegelijk uitvoeren (dual-tasking), namelijk het begrijpen van de vraag die wordt gesteld en het voorbereiden van een respons. In deze dissertatie heb ik geprobeerd om dit stukje van de puzzel op te lossen: aan de hand van een reeks experimenten heb ik onderzocht of repetitie-priming ook voorkomt tijdens gesprekken en of priming dan kan fungeren als ondersteuningsmechanisme. In *hoofdstuk 2* bespreek ik bestaand onderzoek waarin met behulp van taken waarin taal geproduceerd of begrepen moest worden is getest of repetitie-priming ook kan voorkomen in de hierboven genoemde situaties. Eerst heb ik gekeken naar onderzoeken waarbij woorden meteen of na een korte tussenpoos werden herhaald: treedt het repetitie-priming-effect op ongeacht de tijdsspanne tussen de twee herhaalde woorden? Vervolgens wilde ik weten of het presenteren van woorden als onderdeel van een zin een rol speelt bij het al dan niet optreden van het repetitie-priming-effect: is er sprake van hetzelfde priming-effect als mensen eerst het woord begrijpen/produceren als afzonderlijk woord (bijv. *boek*) of als onderdeel van een zin (bijv. *De vrouw liet het boek op tafel liggen*)? Tot slot heb ik de vraag gesteld of het horen van een woord tijdens de voorbereiding op een gesproken respons (bijv. het woord *boek* horen terwijl je je voorbereidt op het uitspreken van het woord *hond*) verdere repetitie-priming belemmert. In de literatuurstudie heb ik een overzicht gemaakt van al het bestaande onderzoek naar het onderwerp en suggesties gedaan voor verder onderzoek. In *hoofdstuk 3* en *4* heb ik me gericht op een van de vragen in *hoofdstuk 2*, namelijk of repetitie-priming afneemt of verdwijnt als mensen naar het te herhalen woord luisteren (het primewoord) terwijl ze zich voorbereiden om iets anders te zeggen. In *hoofdstuk 3* hebben deelnemers geluisterd naar primewoorden (bijv. *radio*) terwijl ze zich moesten voorbereiden op het uitspreken van een ander woord, of geen andere taak hoefden uit te voeren. Vervolgens moesten ze afbeeldingen benoemen waarvoor het zojuist gehoorde primewoord een toepasselijk woord zou zijn (bijv. *radio*) of niet(bijv. *stopcontact*). De resultaten in *hoofdstuk 3* laten zien dat deelnemers afbeeldingen waarvan ze de naam eerder hadden gehoord sneller benoemden dan andere afbeeldingen, ongeacht of ze de woorden voor het eerst hadden gehoord tijdens de voorbereiding op een gesproken respons of wanneer ze geen extra taak hadden. In *hoofdstuk 4* heb ik geprobeerd de resultaten van *hoofdstuk 3* met behulp van EEG verder uit te werken. Daarbij had ik twee belangrijke vragen. Is het EEG-signaal dat wordt uitgelokt door de primewoorden die het brein moet verwerken terwijl het zich voorbereidt op een gesproken respons anders dan dat van woorden die worden verwerkt zonder dat het brein nog extra taken heeft? Is het EEG-signaal dat wordt uitgelokt door herhaalde afbeeldingen anders wanneer deelnemers de bijbehorende naam eerder hebben gehoord terwijl ze zich aan het voorbereiden waren op een respons? Hoewel het experiment geen repetitie-priming-effecten aantoonde, heb ik enkele methodologische aspecten besproken die zouden moeten worden meegenomen in het ontwerp van EEG-experimenten waarin taalbegrip- en productietaken worden gecombineerd. Anders dan in de vorige hoofdstukken - waarin de focus ligt op een combinatie van begrip- en productietaken - richt *hoofdstuk 5* zich uitsluitend op taalbegrip. Hierin heb ik een analyse gemaakt van openbaar beschikbare EEG-data van proefpersonen die aan het luisteren waren naar een audioboek. Door het EEG-signaal voor herhaalde en niet-herhaalde woorden met elkaar te vergelijken, wilde ik uitzoeken of herhaalde woorden priming-effecten uitlokten. Hoewel hier wel enig bewijs voor leek te zijn, verdween het effect op het moment dat ik ook het effect van een ander mechanisme liet meewegen. De resultaten duiden er namelijk voorzichtig op dat de EEG-signalen die werden uitgelokt door de woorden in het audioboek, beter te verklaren zijn op basis van hoe voorspelbaar ze waren gezien de beschikbare informatie in de zin, dan op basis van het feit of het woord al dan niet was herhaald. Kort samengevat is in deze dissertatie onderzocht of repetitie-priming ook kan voorkomen in gesprekken. Allereerst heb ik aangetoond dat wanneer mensen een woord moeten begrijpen terwijl ze zich voorbereiden op een gesproken respons, dit geen effect heeft op repetitie-priming van een begrepen woord. Dit suggereert dat repetitie-priming mogelijk een ondersteunend effect heeft wanneer tijdens gesprekken twee taalgerelateerde taken tegelijk moeten worden uitgevoerd (bijv. wanneer mensen nadenken over een respons terwijl de gesprekspartner nog aan het woord is). Ten tweede heb ik bewijs gevonden dat het effect van repetitie-priming mogelijk beperkt is in natuurlijkere omstandigheden wat comprehensie betreft. Ten derde heb ik enkele methodologische beperkingen uiteengezet omtrent het gebruik van EEG voor het onderzoeken van repetitie-priming en taalkundige dualtasking. Samen maken deze resultaten duidelijk dat er meer onderzoek nodig is om aan te tonen of en hoe repetitie-priming-effecten optreden in dagelijkse gesprekken.

English Summary

Conversations are complex tasks that require the combination of multiple - sometimes overlapping - processes. For instance, when someone asks a simple question, their interlocutor needs to first split the stream of sounds they hear into separate words, and then understand their meaning and their relation to the other words in the question. To correctly answer the question, the interlocutor needs to retrieve a specific set of words from memory, combine them together and articulate them. It is also important that the answer to the question is uttered at the right time: for instance, waiting too long to answer might be taken as a sign of hesitation, interrupting the speaker while they are still talking might be considered as impolite. Given the complex set of processes that conversations involve, why do they seem so effortless? According to an influential model of dialogue, conversations are easy because interlocutors progressively align their representations of the situation under discussion. They do so by repeating parts of their interlocutor's and own utterances (such as words or grammatical structures), a phenomenon called *priming*. Priming of words in particular, i.e., *repetition priming*, has been widely studied in psycholinguistics: in experiments where people are asked to either name a picture with a word or to understand a word, they are usually faster and/or more accurate if the word has been repeated before. While repetition priming appears to be a robust mechanism, it has been rarely tested in situations prevailing in conversation. For instance, many priming studies used single words that immediately repeat each other. Yet, in conversations, this is often not the case: words are usually part of sentences, they can be repeated after some time, and participants may need to dual-task between understanding the incoming question and preparing a response. In this dissertation I attempted to fill this gap: in a series of studies, I tested whether repetition priming can still occur in some of the settings prevailing in conversation and whether it can indeed work as a supporting mechanism.

In *Chapter 2*, I reviewed existing studies that used production and/or comprehension tasks to test whether repetition priming can still occur in the settings mentioned above. First, I looked at studies where words were repeated immediately or after a delay: do repetition priming effects occur regardless of the

distance between two repeated words or not? Second, I asked whether presenting words in a sentence or in isolation matters for the occurrence of priming effects: do people show similar priming effects if they first comprehend/produce the word in isolation (e.g., *book*) or as part of a sentence (e.g., *The woman left the book on the table*)? As a last question, I asked whether hearing a word while preparing a spoken response (e.g., hearing the word *book* while preparing to say the word *dog*) hinders subsequent repetition priming. In the literature review, I summarised the available studies on the topic and suggested avenues for further research.

In *Chapter 3* and *Chapter 4*, I focused on one of the questions reported in *Chapter 2*, namely whether repetition priming vanishes or decreases when people listen to the to-be-repeated word (called *prime word*) while they are preparing to say something else. In *Chapter 3*, participants listened to prime words (e.g., *radio*), while preparing to say another word or without doing any additional tasks. Then, they named pictures (called *target pictures*) that could have the same label of a previous prime word (e.g., *radio*) or a completely different one (e.g., *socket*). The results of *Chapter 3* showed that participants named target pictures whose name they had previously heard faster than non-repeated pictures, regardless of whether they had first heard them while preparing a spoken response or without additional tasks. In the subsequent chapter, *Chapter 4*, I attempted to expand the results of *Chapter 3* using EEG. I asked two main questions. Does the EEG signal elicited by the prime words comprehended while preparing a spoken response differ from that of words comprehended without any additional tasks? Does the EEG signal elicited by repeated *target pictures* differ according to whether participants have previously heard their name while preparing a response or not? While the experiment failed to show repetition priming effects, I discussed methodological aspects to take into consideration for the design of EEG experiments that combine comprehension and production tasks.

Unlike previous experimental chapters that combined production and comprehension tasks, *Chapter 5* focused solely on comprehension. Here, I analysed a publicly available dataset of EEG data recorded while participants listened to an audiobook. I compared the EEG signal of repeated and non repeated words to assess whether repeated words elicited priming effects. While I found some weak evidence of priming, the effect vanished when I also took into account the effect of a different mechanism. More specifically, the data provide some hint at the fact that the EEG signal elicited by the words in the audiobook were better explained by how easily a word could be predicted given the available information in the sentence than by whether the word had been repeated or not.

In sum, this thesis investigated whether repetition priming can still occur in settings prevailing in conversation. First, I have shown that comprehending a word while preparing a spoken response does not affect repetition priming of a comprehended word, which suggests that repetition priming may support conversation during linguistic dual-tasking (for instance, when people plan a response while their interlocutor is still speaking). Second, I have found evidence that, in naturalistic comprehension, the effect of repetition priming may be limited. Third, I have outlined some methodological limitations about the use of EEG to investigate repetition priming and linguistic dual-tasking. Together, these results highlight the need for additional research on whether and how repetition priming effects unfold in every-day conversations.

Acknowledgements

A PhD can be a rewarding but also challenging experience. Completing this work would have been much more burdensome without the help of many amazing people that supported me, professionally and personally.

First, I want to thank my supervisory team for guiding me during this long journey. Antje, you taught me to break complex questions into smaller ones, and to outline hypotheses and ideas in a rigorous way: these skills will always be useful regardless of the career path that awaits me. I also want to thank you for providing support and advice during the uncertain times of COVID-19: it was a complex period and I truly appreciate your help. Suzanne, and Ashley, thank you both for your great supervision. Suzanne, thank you for your enthusiasm and for being always available whenever I knocked on your door to discuss new ideas. Ashley, thank you for taking over and helping me with my MATLAB code and EEG analyses: you have always been available to provide help and guidance and I could not have asked for better support in the last phase of my PhD.

I also want to thank the members of the reading committee Evelien Heijselaar, Herbert Schriefers, and Linda Wheeldon for evaluating this thesis and for providing useful feedback.

None of this work would have been possible without the help from the staff at the MPI, that provided assistance whenever it was needed. Kevin, a special thank you goes to you: you are an amazing Coordinator and I am extremely grateful for all the guidance you provided. I really enjoyed all the friendly chats and the time we spent together on the side activities that we organised/contributed to together. Although I will also mention them later in the acknowledgments, I want to thank here my wonderful paranymphs, who supported me during the last phase of this PhD and with whom I created great memories during these years: I am extremely glad to have shared this path with you.

In addition to the professional and technical advice, I am also grateful to all the people that made this journey a memorable one. I will not make names but I am sure that you will recognise yourselves in the descriptions below (ordered by increasing geographical distance from the MPI).

MPI. At the MPI, I was lucky to be surrounded by a great community of motivating and friendly people. I want to thank the Psychology of Language department for always providing great feedback, and for the time spent together. A special thanks goes to the people with whom I shared afternoon coffees, walks in the city centre, lively discussions in Spanish or Wednesday chats after work: I will always cherish these memories. I also want to thank all the fellow IMPRSers that worked with me on side projects and activities: from the IMPRS conference to the IMPRS Career Event and the D&I working group, collaborating with you was always a pleasure!

Ziekerstraat. There are two flats in Ziekerstraat, almost facing each other, that have become *a home* thanks to two amazing people, with whom I shared the joys and worries that come with a PhD, and much more. From the long working hours during the pandemic to the gym sessions and the improbable meals we cooked together, you made my time in Nijmegen unforgettable. I cannot put into words how much this all meant to me. A warm thank you also to all the people that, during the years, were part of the "Ziekerstraat group", even for a short while: I greatly enjoyed all the time we spent together.

London. This section is for someone that turned 2020 - undoubtedly a complex period - into one the most amazing years so far: thank you for booking that London-Amsterdam flight (and for quarantining almost twenty days for precaution). You have been my safe shore in one of the most challenging periods so far, and you managed to do so from across the Channel. Thank you for always motivating me and believing in me, even when I doubt(ed) myself. You are the best thing that happened to me since I embarked in this PhD adventure: I hope we will go on many more together.

Venice. When I took that *vaporetto* to reach San Servolo in 2013, I did not know that that small island off Saint Mark's Square would become so important. I am extremely grateful to the ones who shared that journey with me, and are still always there for precious advice, proof-reading, and laughs. Thank you for always making the time to meet during our Sunday's Skype calls, be it from New York, St. Petersburg or Venice. A special thanks goes to that BA student who, in 2013, asked endless questions about the course schedule to the *unfriendly girl* with a passion for Linguistics. Almost ten years later, I could not ask for a better shoulder to rely on: I hope I will be able to support you during your PhD as you supported me during mine.

Prato. No matter where I move, I know that there is always a small bunch of people that awaits me when I go back to my home town. Our nights together

during the holidays were a much needed distraction from my PhD: thank you for all the chats and dinners together. This section would be incomplete without thanking my best friend: you have been there at all stages of this PhD and beyond, and supported all my choices. Thank you for always reminding me where I come from: you are "*the color of the wheat fields*".

Prato. Ovunque io vada, so che ci sarà sempre un piccolo gruppo di persone ad aspettarmi quando torno nella mia città. Le serate passate assieme durante le vacanze mi hanno permesso di rilassarmi e distrarmi dal mio dottorato: grazie per tutte le chiacchiere e cene assieme. Questa sezione sarebbe incompleta se non ringraziassi la mia migliore amica: sei stata presente in ogni fase di questo dottorato (e oltre), e hai supportato ogni mia scelta. Grazie per ricordarmi sempre da dove io venga: sei *"il colore del grano"*.

My parents. You wanted me to become a lawyer but I followed quite a different path. Yet, you never second-guessed my choices, and you always supported me as best as you could. You have worked and sacrificed much so that I could study. Mum, thank you for always encouraging me to do better. Dad, thank you for listening to me on that day in 2012 and for convincing me to move to Venice. I hope I will be able to repay all your efforts.

I miei genitori. Avreste voluto che diventassi un avvocato ma ho preso una strada totalmente diversa. Eppure non avete mai messo in dubbio le mie scelte, e mi avete supportata come meglio potevate. Avete lavorato e sacrificato tanto perché io potessi studiare. Mamma, grazie per avermi sempre spronata a studiare e a fare meglio. Babbo, grazie per avermi ascoltato in quel lontano autunno del 2012 e per avermi incoraggiata a trasferirmi a Venezia. Spero di poter ripagare tutti i vostri sforzi.

Curriculum Vitae

Federica Bartolozzi was born in Prato (Italy) in 1993. In 2015, she obtained a bachelor's degree in Languages, Linguistics, and Cultural Studies from Ca' Foscari University of Venice (Italy). After that, she moved to the United Kingdom, where she completed her master's degree in Psychology of Language at the University of Edinburgh in 2016. In September 2017, she started her Ph.D. in the Psychology Department of the Max Planck Institute for Psycholinguistics.

Publications

Bartolozzi, F., Jongman, S. R., & Meyer, A. S. (2021). Concurrent speech planning does not eliminate repetition priming from spoken words: Evidence from linguistic dual-tasking. Journal of Experimental Psychology: Learning, Memory, and Cognition, 47(3), 466.

Nieuwland, M. S., Barr, D. J., Bartolozzi, F., Busch-Moreno, S., Darley, E., Donaldson, D. I., ... & Von Grebmer Zu Wolfsthurn, S. (2020). Dissociable effects of prediction and integration during language comprehension: Evidence from a large-scale study using brain potentials. Philosophical Transactions of the Royal Society B, 375(1791), 20180522.

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., ... & Huettig, F. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. ELife, 7, e33468.