MAX PLANCK INSTITUTE
FOR PSYCHOLINGUISTICS

WWW.MPI.NL

MEMORY FOR SPEAKING AND LISTENING

Eirini Zormpa

Memory for speaking and listening

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.

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, Munich, Germany.

ISBN: 978-94-92910-19-6

Printed and bound by Ipskamp Drukkers, b.v.

Memory for speaking and listening

Proefschrift

ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,

volgens besluit van het college van de decanen

in het openbaar te verdedigen op vrijdag 18 december 2020

om 13:30 uur precies

door

Eirini Zormpa

geboren op 1 december 1991

te Serres (Griekenland)

Promotor:

Prof. dr. Antje S. Meyer

Copromotoren:

Dr. Laurel E. Brehm (Max Planck Institute for Psycholinguistics)

Dr. Renske S. Hoedemaker (Max Planck Institute for Psycholinguistics)

Manuscriptcommissie:

Prof. dr. Herbert J. Schriefers

Prof. dr. Eva Belke (Ruhr-Universität Bochum, Duitsland)

Dr. Andreas Mädebach (Universitat Pompeu Fabra, Spanje)

Table of Contents

1 | General introduction

Memory affects all aspects of our cognition. That might be the reason why "memory" appears in the title of 1 out of 15 papers in psychology, more frequently even than "thinking" (Simonsohn, 2013). Memory also underlies our ability to use language to communicate with others. Firstly, memory stores the content we want to express, like our world knowledge and experiences, as well as the tools we need to verbally express ourselves. Secondly, smooth conversation requires us to keep track of what was said and by whom, which requires memory.

The relationship between language and memory is not unidirectional: Language can affect memory, too. Firstly, the attributes of a word affect how well we remember it. For instance, we remember real words much better than non-words. Other lexical attributes, like a word's length, frequency, and familiarity, are also good predictors of memory for that word. Secondly, the way we choose to use language affects how well we remember the words we say. For instance, we remember the names of objects in our environment better if we mention them in our speech than if we do not and even the volume and manner of our production can play a role for memory.

Given these links, it is quite surprising that language and memory researchers typically do their work rather independently of each other. There is a journal called *Journal of Memory and Language*, but it could just as well be called Journal of Memory OR Language, as most papers concern only one of these aspects of cognition. Even in those studies that do consider both, the relationship between memory and communication in particular is rarely researched. This is problematic because the most common way in which we use language is to communicate with each other.

In this thesis, I was interested in how well people remember what was said in conversation. The joint investigation of memory and conversation is important because we learn a lot from talking to others: we learn new words, new facts, and, in certain points of our lives, even whole languages. As such, getting better insights into the factors that influence what we retain from conversations can have important implications for learning more generally. My aim in this thesis was to bring together research on language and research on memory to gain a better understanding of the processes affecting people's memory for conversation.

Studying communication is important, but it is also difficult: Communication is messy and gives little opportunity for experimental control. As a result, most psycholinguistic studies tend to use language that is more fluent and controlled than the language used in everyday communication. Here, I have tried to balance the desire to study naturalistic language use with the need for experimental control by taking an incremental approach. I started by looking at language processing on the level of the individual, testing language production and language comprehension separately before combining them in conversation. In this way, I was able to study effects in a tightly controlled experimental environment and then test how one of these effects played out in a conversational context. I focused entirely on the level of words, which in this case were always concrete nouns. I focused on words for two reasons: First, words are the building blocks of longer language segments and concrete nouns often carry a lot of meaning in a sentence. As such, memory for those words is indicative of memory for a sentence. Secondly, words are easier to test in a controlled manner than sentences are.

In the remainder of this chapter, I introduce some concepts that will be important for following this thesis. First, I discuss how known memory phenomena can map onto word production processes to show which of those processes might enhance memory for words. Then, I discuss work on how what is emphasised in an utterance influences memory. Finally, I bridge these two lines of research and discuss how speakers' actions and interests can influence their own and their interlocutors' memory for their conversation.

Memory phenomena

The first step towards identifying what factors affect memory for conversations is to establish what factors affect memory for words in general. Memory research has often been concerned with the question of why certain items are remembered better than others presented in the same list. For example, this work has often assessed whether a process gives a relative benefit to an item or instead, a relative handicap to that item's competitors (other items in the list). Another common question is whether an item itself receives a benefit from a process or if the connection between that item and its neighbours receives a benefit. In this thesis, I am interested in memory for the items themselves, so I will limit the introduction to item-specific processing. Two phenomena in particular, the generation effect and the production effect, are discussed in more detail because they were studied in the experiments described in Chapters 2 and 3.

The generation effect

The generation effect refers to the finding that we remember words we come up with ourselves better than the words that we read (Bertsch, Pesta, Wiscott, & McDaniel, 2007; Slamecka & Graf, 1978). Crucially for dialogue, the generation effect has also been replicated using auditory stimuli: We remember words that we come up with ourselves better than words that we hear (Dew & Mulligan, 2008). For conversation, this would imply that we remember our own contributions better than those of our conversational partners.

In tests of the generation effect, participants first complete a study phase in which they see words in a generate condition or in a read condition and then complete a memory task. In the generate condition, participants usually see a cue word and the first letter of a target word which is in some way related to the cue word (e.g. "sheep" - "g _ _ _"). They then need to come up with the correct word and say or write both words (e.g. "sheep" - "goat"). In the read condition, participants see both the cue and the target (e.g. "sheep" - "goat") and have to say or write both words. The cue-target relationship tested in the example is that of semantic associate. Other relationships that lead to a generation effect include synonyms, antonyms, rhymes, translations, and definitions.

The memory task, which is when the generation effect is observed, can occur immediately after, a few minutes after, or a day or more after the study phase. The memory task can also take different forms. The generation effect has been replicated with recognition memory tasks and with recall tasks. In recognition memory tasks, the main method used in this thesis, participants see the same words as in the study phase along with an equal number of new words and decide which words they saw in the first phase. In recall tasks, participants write down as many words they read or generated as they can.

Different accounts have been proposed for the generation effect (Bertsch et al., 2007). A well supported account is the *distinctiveness account* (Hunt & McDaniel, 1993; Hunt & Worthen, 2006), according to which generated items contain unusual (or distinct) features that can be used as a heuristic at test to aid recognition. Other influential accounts include the *item-order account* (McDaniel & Bugg, 2008), according to which item-specific processing is enhanced at the expense of relational processing, such as order processing.

The production effect

The production effect refers to the finding that we remember the words we say aloud better than the words we say silently in our heads (MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010). For instance, participants have been found to recognise the words that they read aloud 10-20% better than the words they read silently (MacLeod et al., 2010). Like the generation effect, the implication of the production effect for dialogue is that we should remember our own speech better than an interlocutor's.

Studies of the production effect are similar to those of the generation effect in that they consist of a study phase and a memory task. In the study phase, participants see words that they have to read aloud or silently, depending, for instance, on the font colour of the word. The memory task can, again, take place from immediately after the study phase to a week after the study phase. Like the generation effect, the production effect has been replicated in memory tasks testing both recognition and recall.

The production effect also applies to words we mouth, whisper, shout, sing, or write, though the size of that effect varies depending on mode of production (Forrin, MacLeod, & Ozubko, 2012; Mama & Icht, 2016; Quinlan & Taylor, 2013). This effect is also observed when comparing only hearing a word to hearing a word and then saying it aloud (Mama & Icht, 2016). A finding even more relevant to communication is that, when the production effect was tested with pairs of participants, people remembered what they had said themselves better than what their partner had said (MacLeod, 2011).

Similar to the generation effect, the production effect has been explained most successfully under the distinctiveness account. That is, the act of saying a word aloud is thought to add a distinctive trace that can be used heuristically at test (Dodson & Schacter, 2001; MacLeod et al., 2010; Ozubko, Major, & MacLeod, 2014). The production effect is of obvious relevance for memory of conversational content, as it should yield a memory advantage for the speaker's own utterances compared to the partner's utterances.

Word production

The memory literature has established that both coming up with words and saying words aloud are robust ways of enhancing word memory with long-lasting effects. However, this research has largely disregarded the linguistic processes whereby words are generated and produced. A key concern of my thesis was to gain further insight into the basis of the generation effect and the production effect in processes of word production[1]. That is, I was interested in which linguistic processing components are engaged when people say, rather than listen to, words and in how the engagement of these components could support later memory for the word. To address these issues, a working model of word production is required.

[1]In the psycholinguistic literature, *language production* or *word production* encompasses the entire process from generating the message to be expressed all the way to the auditory (or other) signal, not just the process of speech production implied in the production effect.

In this thesis, I follow the model described in Levelt, Roelofs, and Meyer (1999). Much of the research this model is based on used picture naming studies (like the ones used in this thesis), so perhaps an illustration of how word production occurs in these studies is useful.

Say that you are confronted with the picture of a baby goat and asked to name it. First, you will look at the picture and process its visual features. Then, you will recognise the animal depicted in the picture and activate a pre-verbal concept of it. Up until this point, you have not used any language. Now, however, you need to select a lexical concept to name the picture (conceptual preparation). The basic-level name for the picture is, of course, "goat" but you could also call it "animal" or "kid[2]", or if you are familiar with that particular baby goat, you could call her "Sally". All of these concepts receive some activation. Which concept is eventually selected depends on the situation in which you are naming the picture. In addition to these concepts, related concepts, e.g. "sheep", also receive some activation. After you select an appropriate concept, you need to activate the corresponding lemma (lexical selection). Lemmas are amodal representations which contain syntactic information, like grammatical gender. In the next stage, the corresponding phonemes are selected and combined into syllables (form encoding). Finally, a gestural plan is created (phonetic encoding) and executed (articulation) to produce sound.

Relationship between memory and word production processes

It is currently unclear which of these word production processes are affected by the generation effect and the production effect, introduced previously. Establishing which processes underlie these memory phenom-

[2] The word *kid* can also be used to refer to the young of goats. In fact, that was the word's original meaning, which was extended to refer to human children in the late 16th century.

ena would provide a way to explain some of the factors influencing memory during day-to-day communication that is grounded in psycholinguistic research. Much of the work on these effects, especially the generation effect, has been in the written modality, making frequent use of orthography and phonology. However, to be able to link processes of word production to memory effects, it is important to first replicate these memory phenomena using tasks that more closely resemble conceptually mediated language production, and hence conversation. In my thesis, I studied the generation effect and the production effect using picture naming tasks.

In Chapter 2, I used an adaptation of the picture naming task to ensure that the generation effect and the production effect obtain in a conceptually mediated language production task. Participants saw pictures either with the picture names or with unreadable labels superimposed and had to name these pictures either aloud or silently. They were more accurate when recognising pictures they had named themselves (unreadable label condition) rather than read, and pictures they had named aloud rather than silently.

In Chapter 3, I followed up on the findings of Chapter 2 and tested if naming a picture improves memory not only for the picture, but also for the picture name itself. My aim in this chapter was to ensure that generating and producing a word aloud affects linguistic representations. This was a necessary step for the following chapters, as my interest in this thesis was in the linguistic content that people remember from conversations. In addition to the effects of generation and oral production, I tested the effect of processing time on memory. Processing time was measured in different ways, including naming latencies (time taken to name a picture) and gaze durations (overall looking times to each picture). Looking at different processing times measures helped me disambiguate between different processes and evaluate their contribution to the observed memory benefit.

Memory for conversation

Consistent with the generation and production effects, recent psycholinguistic research on conversation and memory has reported a memory advantage for speakers (Fischer, Schult, & Steffens, 2015; McKinley, Brown-Schmidt, & Benjamin, 2017; Yoon, Benjamin, & Brown-Schmidt, 2016). That is, people remember what they have said themselves better than what they have heard from their interlocutors. This is somewhat surprising, considering that, intuitively, conversations seem to be about communicating with others. Establishing whether people do remember their own utterances better than their interlocutors' is an important step towards understanding the dynamics of communication. Before turning to communication, however, we first need to consider some key properties of language comprehension.

A critical notion in the discussion of memory for comprehended language is that of *focus*. Information that is in focus is considered to be the most important in a sentence (Chomsky, 1971). For example, in "It is carrots that Sally likes", "carrots" is the most important part of the sentence. Focused items are considered to attract attention which confers processing benefits to them (Foraker & McElree, 2007). Importantly for this thesis, one such benefit is that focused items are remembered better than unfocused or neutral items (e.g., Fraundorf, Watson, & Benjamin, 2010; Sturt, Sanford, Stewart, & Dawydiak, 2004). However, the methods used to study focus often bear little resemblance to everyday communication. The tasks are often in the written modality and usually feature clefts as a way to manipulate focus. Clefts, like in the example above, are extremely rare, appearing in less than 0.1% of English sentences (Roland, Dick, & Elman, 2007).

In Chapter 4 I used question-answer pairs, which are quite common in communication, to study the effect of focus on memory for conversations. In this study, participants listened to recorded question-answers pairs, and were found to remember focused information (answers) better than non-focused information (questions).

The study in Chapter 5 used a communicative task in which participants took turns asking and answering each other questions. In question-answer pairs, questions are produced (which confers a memory benefit) and answers are focused (which also confers a memory benefit). This chapter thus brings together the two lines of research present in this thesis to ask how language production processes and focus affect memory in conversation.

Combined, these studies highlight the interdisciplinary nature of the thesis which brings together research on the fields of memory and language. The ultimate aim is to develop a better understanding of the factors that underlie the memory asymmetries between language production and language comprehension during communication.

2 | The production effect and the generation effect improve memory in picture naming[1]

Abstract

The production effect (better memory for words read aloud than words read silently) and the picture superiority effect (better memory for pictures than words) both improve item memory in a picture naming task (Fawcett, Quinlan, & Taylor, 2012). Because picture naming requires coming up with an appropriate label, the generation effect (better memory for generated than read words) may contribute to the latter finding. In two forced-choice memory experiments, we tested the role of generation in a picture naming task on later recognition memory. In Experiment 1, participants named pictures silently or aloud with the correct name or an unreadable label superimposed. We observed a generation effect, a production effect, and an interaction between the two. In Experiment 2, unreliable labels were included to ensure full picture processing in all conditions. In this experiment, we observed a production and a generation effect but no interaction, implying the effects are dissociable. This research demonstrates the separable roles of generation and production in picture naming and their impact on memory. As such, it informs the link between memory and language production and has implications for memory asymmetries between language production and comprehension.

[1]Adapted from 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. doi: 10.1080/09658211.2018.1510966.

Introduction

Over the decades, memory research has identified a set of encoding strategies that can enhance retention even after a single exposure to an item. One of these is simply producing a name aloud: reading a word aloud, as opposed to silently, improves recognition memory by 10 to 20%. This is known as the *production effect* (e.g. Conway & Gathercole, 1987; MacLeod & Bodner, 2017; MacLeod et al., 2010). Although usually tested with words, this effect has also been reported in a picture naming task (Fawcett et al., 2012). Additionally, the production effect also arises across different manners of production, such as singing, whispering, mouthing, and typing (Forrin et al., 2012; Quinlan & Taylor, 2013). This effect has been most successfully explained under the *distinctiveness account*. The distinctiveness account posits that producing a name aloud adds an extra feature, that of the production record, to an item's memory trace (Conway & Gathercole, 1987; MacLeod et al., 2010; Ozubko & MacLeod, 2010), which can be used heuristically at test to improve memory performance (Dodson & Schacter, 2001). Two further manipulations of encoding condition that impact memory performance are, first, studying items as pictures versus as words (the *picture superiority effect*) and, second, generating labels for items versus reading existing labels (the *generation effect*). In this article, we examine how the production effect, the generation effect, and the picture superiority effect interrelate. We extend the claim that producing a picture name aloud enhances memory through speaking to propose that it also enhances memory through the active generation of a label. We test this claim in two recognition memory experiments where we dissociate the role of overt word production from the role of label generation in item memory.

In order to account for the effects of word production on picture naming, we first consider how pictures themselves influence memory compared to words. Picture stimuli tend to be remembered better than words, a finding known as the picture superiority effect (Paivio, Rogers, & Smythe, 1968). This effect is robust to a variety of manipulations, including modality changes between study and test. That is, memory for items studied as pictures is better than memory for items studied as words, even when memory is tested with words (Borges, Stepnowsky, & Holt, 1977). Distinctiveness has also been successful in accounting for the picture superiority effect (Curran & Doyle, 2011). Pictures are thought to be remembered better than words because the visual representations evoked by pictures are more distinctive than those evoked by printed words (Nelson, Reed, & McEvoy, 1977). It has also been argued that pictures undergo more extensive conceptual processing than words (Weldon & Roediger, 1987). This suggests that the picture superiority effect and the production effect may have a common basis: Both effects rely on distinctiveness of memory traces as a result of rich encoding at study.

The question of how the picture superiority effect and the production effect jointly relate to item memory was tested by Fawcett et al. (2012). They hypothesised that if the picture superiority effect and the production effect both rely on distinctiveness, the two effects should interact to give pictures named aloud an especially distinctive memory trace. To test this hypothesis, Fawcett and colleagues provided participants with picture and word stimuli at study and asked them to mouth the item names, producing speech movements but no overt speech, or to generate the item names using inner speech only. In a Yes/No recognition task for the pictures and words, participants showed increased sensitivity for mouthed than for internally named items, showing a production effect, and increased sensitivity for pictures than for words, showing a picture superiority effect. The predicted interaction also occurred, such that mouthing benefited items more in the picture condition than in the word condition. Fawcett and colleagues interpreted this interaction as evidence that distinctiveness underlies both the production effect and the picture superiority effect.

In Fawcett and colleagues' study, the distinctiveness of pictures was attributed to their visual characteristics. However, note that when words and pictures are named - be it overtly or covertly - the processes mediating between the visual input and the linguistic representations of the names also differ. Picture naming is conceptually driven: Speakers need to identify the concept represented in the picture, select an appropriate lexical unit, and retrieve the corresponding phonological and articulatory commands (Dell & O'Seaghdha, 1992; Indefrey & Levelt, 2004; Levelt et al., 1999). By contrast, reading primarily involves the mapping of orthographic onto phonological and articulatory representations. This process typically entails the activation of semantic and conceptual representations, but it does not hinge on these processes, as evidenced by the fact that skilled readers can readily read non-words (Rosson, 1983; Theios & Muise, 1977). Thus, picture naming and word reading rely on different cognitive processes and these processes may affect item memory differentially.

The hypothesis that the differences between reading and picture naming in post-perceptual processes may contribute to differences in recognition memory is further motivated by another observation from the memory literature, the generation effect (e.g. Bertsch et al., 2007; Slamecka & Graf, 1978) This is the finding that participants are better at remembering words that they generated themselves than words that they read. In a common version of the paradigm used to elicit this effect, participants either read antonym pairs (e.g., "hot – cold") or are presented with the first member of the pair and generate the antonym themselves (e.g., "hot - c_ _ _"). Other ways to elicit the generation effect include the generation of words from synonyms (Slamecka & Graf, 1978), semantic associates (Begg, Snider, Foley, & Goddard, 1989), translations (Slamecka & Katsaiti, 1987), rhymes (Slamecka & Graf, 1978), and definitions (Forrin, Jonker, & MacLeod, 2014; MacLeod et al., 2010). In all of these cases, words that are self-generated are remembered better than words that are just read. Importantly, most of these generation rules require semantic processing. By contrast, no generation effect is observed for non-words generated using either rhyme or letter transposition rules (Nairne, Pusen, & Widner, 1985; Nairne & Widner, 1987), both of which are known to lead to a generation effect for words. This pattern suggests that the generation effect stems from the conceptual or semantic processing of the stimuli at study. The exact mechanism behind the generation effect has not been settled in the literature, but similar to the production effect and the picture superiority effect, distinctiveness has been proposed as a likely explanation (Gardiner & Hampton, 1985, 1988; Kinoshita, 1989).

Since picture naming relies more on conceptual processing than word reading does, this raises the question of whether part of the picture superiority effect observed in previous work may be attributable to a generation effect, arising during the conceptual and linguistic encoding rather than the visual encoding of the stimuli (for similar proposals see Smith & Magee, 1980; Weldon & Roediger, 1987). Therefore, it is crucial to dissociate which processes contribute to the memory benefit for named pictures.

Beyond its impact on fundamental memory research, discovering the role of the generation effect and the production effect in picture naming also provides valuable insights for psycholinguistics. Recent work has reported memory asymmetries as a function of speaking versus listening. More specifically, language production[2] seems to elicit a memory benefit, such that speakers have superior item memory compared to listeners for items referred to in a conversational context (Hoedemaker, Ernst, Meyer, & Belke, 2017; McKinley et al., 2017; Yoon et al., 2016). Disassociating the roles of the generation and production effects, both of which relate to mechanisms of language production, allows us to characterise the asymmetries between language production and comprehension and, more broadly, provides insight on how language and memory intertwine.

Current Study

The main goal of the present study was to examine whether the generation of picture names would improve recognition memory for the pictures via a more distinctive conceptual and linguistic representation of the stimuli. To ensure that generation rather than visual distinctiveness was the factor influencing performance, participants were presented with pictures in all conditions of our study. In one condition (picture+word condition), picture names were provided as labels superimposed on the pictures, whereas in the other condition (picture-only condition), labels were replaced by random patterns of equal visual complexity requiring participants to generate the names themselves (see also Weldon & Roediger, 1987). This allowed us to compare a picture-only to a picture+word condition rather than comparing a picture to a word condition, making the two stimulus types well matched with respect to visual distinctiveness. The prediction was that the presence of the correct labels should greatly facilitate the generation of the object names in the picture+word condition. One may think of the correct labels as identity primes for the picture names. If the conceptually driven generation of object names facilitates later recognition of the objects, performance in a recognition test should be better in the picture-only than in the picture+word condition.

Following Fawcett et al. (2012), we also manipulated whether the pictures were overtly named or not, aiming to elicit a production effect. In the silent condition, participants produced the object names in inner speech, and in the aloud condition, they produced the object names aloud. In the lat-

[2] Note that the production effect refers to the comparison of overt versus inner speech. Language production refers to the entire process of producing language starting with the conceptual processing of what is being communicated and culminating in articulation.

ter condition, we elected to use overt rather than mouthed speech, unlike Fawcett et al. (2012), because we wanted to measure the participants' response latencies. This meant that the effects of the engagement of articulatory gestures and auditory feedback were confounded, but separating them was not of interest in the current study. For evidence that audition improves memory see MacLeod (2011) and Forrin and MacLeod (2018).

The study design followed Fawcett and colleagues' Experiment 3. The four conditions created by crossing stimulus and response type were intermixed throughout the experiment, with visual cues (coloured frames around the pictures) indicating whether the picture was to be named aloud or silently. Deviating from the earlier study, items were presented in a delayed naming paradigm (discussed in the Method section of Experiment 1) in order to encourage participants to generate object names in both the aloud and the silent conditions. After the naming task, participants completed a 20-minute filler test—a computerised version of Raven's Advanced Progressive Matrices (Raven, 1998). The final experimental task was a self-paced Yes/No recognition memory test.

In Yes/No recognition tests, participants' tendency to respond positively or negatively can influence hit rates. This confound is avoided by computing each person's sensitivity and bias in line with signal detection theory (Macmillan & Creelman, 2005). Here we use an application of signal detection theory to mixed-effect models using log-odds ratios (DeCarlo, 1998; Wright, Horry, & Skagerberg, 2009). The benefit of this approach is that it accounts for participant and item variability. As with any other signal detection analysis, the primary dependent measure was sensitivity. We predicted higher sensitivity for items in the picture-only condition than in the picture+word condition, reflecting the generation effect. We also predicted higher sensitivity in the aloud condition than in the silent condition, reflecting the production effect. Assuming that the generation and production effects stem from different aspects of the language production system, we predicted no interaction between stimulus type and response type. This is consistent with the findings of Experiment 1 of Forrin et al. (2014), in which the contributions of the generation and production effects on recognition memory for words were found to be independent.

Experiment 1

This experiment was designed to test the role of the generation effect and the production effect in recognition memory for items studied in a picture naming task. It consisted of three tasks run in one session of approximately one hour. Participants first completed a picture naming task (the study phase) where items varied in their stimulus type (whether the name was primed by the label or not) and in the type of response they elicited (aloud or silent). Participants then completed a filler task, and finally a forced-choice recognition memory task (the test phase).

Method

Participants

Forty-three native Dutch speakers (34 female, age range: 18-30 years) participated in the experiment. They were recruited from the Max Planck Institute for Psycholinguistics participant database and received 10 € each. All participants had normal colour vision. One participant was excluded due to experimenter error and another due to slow response times, leaving 41 participants in the analysis. An a priori power calculation using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) indicated that with 32 items per condition this sample was sufficient to detect effects of at least Cohen's d = 0.48 (a medium effect size) with 95% power[3]. Ethical approval to conduct the study was given by the Ethics Board of the Social Sciences Faculty of Radboud University.

Materials and Design

The stimuli were 256 pictures selected from the BOSS photo database, presented in 400x400 pixel resolution (Brodeur, Dionne-Dostie, Montreuil, & Lepage, 2010; Brodeur, Guérard, & Bouras, 2014). Because the norming data from the BOSS database refer to the English names of the objects, name agreement scores were obtained from eight Dutch native speakers who did

[3]In G*Power, the test family selected was "F tests", the statistical test was "ANOVA: Repeated measures, between factors", and the type of power analysis was "Sensitivity: Compute required effect size – given α and sample size". The number of groups (corresponding here to naming conditions) was 4 and the number of measurements (corresponding here to the number of items in the naming task) was 128. This approximation takes into account the limitations of G*Power, in which computing power for linear mixed effects models or repeated measures ANOVA with more than one within-subjects factor is not supported.

not participate in the experiments. Frequency scores for Dutch were obtained from the SUBTLEX-NL corpus (Keuleers, Brysbaert, & New, 2010). The 256 items were divided into two sets (A and B), which were matched on name agreement (M_A = 0.81, M_B = 0.71; $t(254)$ = -0.18, p = 0.40), $\log_{10}$ word frequency (M_A = 2.16, M_B = 2.20, $t(254)$ = -0.33, p = 0.75), familiarity (M_A = 4.35, M_B = 4.38, $t(254)$ = -0.58, p = 0.26), visual complexity (M_A = 2.36, M_B = 2.34, $t(254)$ = 0.33, p = 0.70), manipulability (M_A = 2.79, M_B = 2.88, $t(254)$ = -0.84, p = 0.19), and length (in letters) (M_A = 6.70, M_B = 7.03, $t(254)$ = -0.93, p = 0.73). A list of the pictures used can be found in Appendix A.

Eight lists were constructed using the two picture sets. In four lists, pictures from set A were used as targets: They were named at study and presented again at test as old items, and set B pictures were presented as foils at test only. In the four remaining lists, set B pictures were used as targets and set A pictures as foils.

At study, participants saw each of the 128 items in one of the four possible naming conditions. These naming conditions were created by crossing stimulus type and response type (picture-only aloud, picture-only silent, picture+word aloud, and picture+word silent). There were 32 items per condition; assignment of items to conditions was counterbalanced across lists. In the picture+word condition, the names of the objects were presented as printed labels in 20 pt Arial font superimposed upon the pictures. In the picture-only condition, the printed labels were replaced by random patterns of matched visual complexity. These patterns were created from non-words which were formed by combining the ending of a target word to the beginning of a different target word. Each letter of the non-words was divided into segments (3x3 grid), which were then assigned random locations and orientations. To indicate the silent versus aloud conditions, coloured frames were presented around the items: green (RGB: 0, 255, 0) for aloud production and red (RGB: 255, 0, 0) for silent production (see Figure 2.1 for stimuli examples).

At test, all 256 pictures were included, with probe type (target versus foil) manipulated between items. The 128 studied pictures served as targets and the 128 non-studied pictures as foils. Target pictures appeared as they had at study, i.e. with or without superimposed printed names and with a green or a red frame. Foil pictures also appeared with a green or red frame around them and a superimposed readable or unreadable label, as if they had belonged to a naming condition. This enabled us to calculate separate false

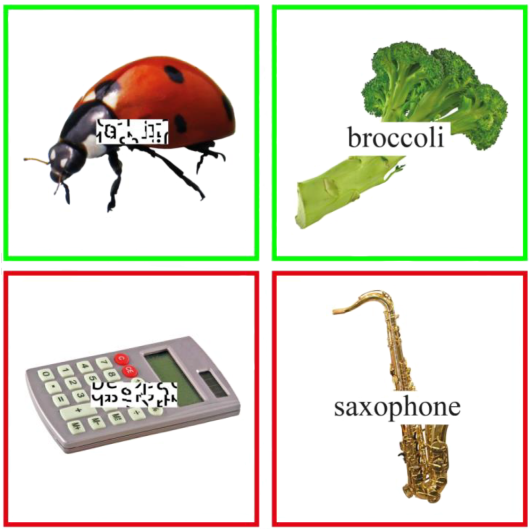


Figure 2.1: Examples of the stimuli used in each naming condition. Top left: picture-only aloud, top right: picture+word aloud, bottom left: picture-only silent, bottom right: picture+word silent. Green frames signalled aloud trials and red frames signalled silent trials.

alarm rates for each condition (see also Fawcett et al., 2012; Forrin, Groot, & MacLeod, 2016).

Procedure

The experiment was controlled by Presentation (version 18.3; Neurobehavioral Systems, Berkeley, CA, USA) and displayed on a 24" monitor (1920x1080 pixel resolution).

Picture naming. Trials started with a fixation cross presented in the middle of the screen for 500 ms, followed by a blank screen for 500 ms. Next, a picture appeared in the middle of the screen. After 800 ms a green or red frame appeared around the picture, and both picture and frame remained on screen for 1200 ms. A blank screen appeared for 500 ms, serving as an inter-trial interval (see Figure 2.2 for an example of the sequence of events in this task). Participants were instructed to start preparing the picture name as soon as they saw the picture, but to produce it only after seeing the frame. As described in the Introduction, a delayed naming paradigm was used to make sure that the participants generated the picture names both in the silent and the aloud condition.

The experiment was preceded by a practice phase of 12 trials, in which 12 pictures that did not appear in the main experiment were used. Three pictures appeared in each of the naming conditions. The practice trials had the same structure as experimental trials with the exception that the reaction time for each aloud trial was displayed on the screen before the next trial began to encourage participants to respond as fast as possible.

Filler task. The filler task was a computerised version of Raven's Advanced Progressive Matrices (Raven, 1998). It consisted of 36 multiple-choice problems that become progressively more difficult. In each trial, participants saw a pattern which needed to be completed and eight possible solutions. Participants had 20 minutes to solve as many problems as possible.

Recognition memory. Memory was tested in a self-paced Yes/No recognition memory task. Participants saw one picture at a time and were instructed to press the left arrow key if they had named the picture during study and the right arrow key if they had not done so. As a reminder, "JA" (yes) was printed at the bottom left corner of the screen and "NEE" (no) at the bottom right. Trial order was randomized within Presentation.

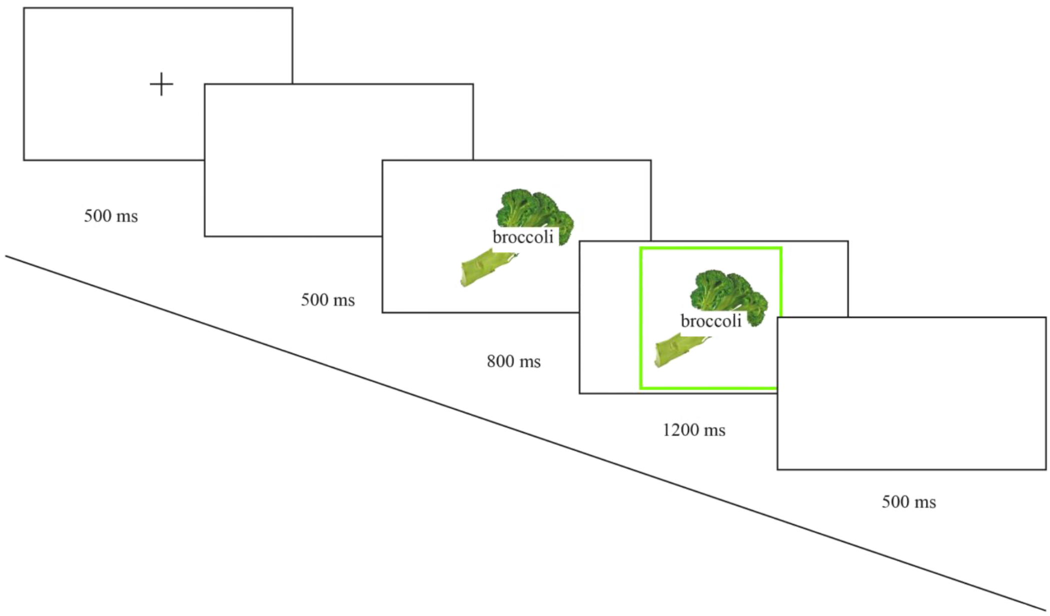


Figure 2.2: Example of the sequence of events during a picture naming trial. This is a picture-only aloud condition.

Analysis

Memory performance was the primary dependent variable in this study. Naming data were used only to exclude trials.

Naming. The naming task was designed to encourage participants to prepare to name all the pictures they saw, even those in the silent naming condition. To test the effectiveness of this design we ran a norming experiment to estimate the average naming times for our stimuli and contrast them with the average naming times in the main experiment. In the norming experiment, eight Dutch native speakers who did not participate in the main experiment named aloud all the pictures. Their responses were annotated in Praat (Boersma & Weenink, 2016) by the first author and their naming latencies were extracted. In the picture-only condition, norming participants started speaking 1000 ms (*SD* 456 ms) after picture onset. In contrast, participants in the main study started speaking 477 ms (*SD* 185) after the frame (the response cue) was presented. In the picture+word condition, norming participants started speaking 705 ms (*SD* 367 ms) after picture onset and main study participants started speaking 402 ms (*SD* 84 ms) after the frame was presented. This demonstrates that in both overt conditions, participants began planning as soon as they saw the image, allowing them to respond faster than the norming participants when given the cue to produce a response. Therefore, we concluded that our participants covertly named the pictures in the silent conditions. One individual with unusually slow naming times (*M*=572 ms) was excluded, as mentioned in the Participants section.

Memory. Responses were modelled as a function of probe type (target vs. foil), stimulus type (picture-only vs. picture+word), response type (aloud vs. silent), and their interactions (Wright et al., 2009). This provides measures of bias (how likely participants were to say "Yes" to any item) and sensitivity (how likely participants were to say "Yes" to an old item), and the way that these measures were modulated by the predictors; as such the analysis is conceptually similar to the standard signal detection analysis and is consistent with current best practices (e.g. Fawcett & Ozubko, 2016; Fraundorf et al., 2010; Jacobs, Dell, Benjamin, & Bannard, 2016; McKinley et al., 2017; Yoon et al., 2016).

Memory performance was analysed using a logistic mixed effects model that was run using the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in R (version 3.4.0 R Core Team, 2017) with the optimiser *BOBYQA* (Powell, 2009). All coefficients represent changes in log-odds of "Yes" versus "No" responses as a function of the predictor. "No" responses were coded as 0 and "Yes" responses as 1. All fixed effects were sum-to-zero coded with foils, picture+word and silent conditions coded to -0.5 and targets, picture-only, and aloud conditions to 0.5. The model with the maximal random effects structure included random intercepts for participants and items, random slopes for the effects of stimulus type, response type and their interaction on items, and random slopes for the effects of probe type, stimulus type, response type and their interactions on participants. This model failed to converge, so the interaction terms for the by-subject and by-item random slopes were removed. In this model, the random slopes for stimulus type and response type effects on items were highly correlated with the random intercept for items (1.00 in both cases), indicating overfitting, and so both random slopes were excluded. Furthermore, the by-item random slope for response type was removed because it explained very little variation (0.05). The final model included random intercepts for participants and items, as well as random slopes by subject of probe type and stimulus type. Reported p-values were obtained using likelihood ratio tests comparing the full model to a model with the same random effects structure but without the fixed effect in question, which has been shown to be more conservative than Wald z tests for small samples (Agresti, 2007). Reported 95% confidence intervals were obtained using the *profile* method of the *confint* function.

Results

Accuracy in the memory task was high (90%). As Figure 2.3 shows, the hit rates were higher in the picture-only than in the picture+word condition, and they were higher in the aloud than in the silent naming condition. In contrast to the hit rates, the false alarm rates were minimally affected by the experimental conditions. This meant that, in a signal detection analysis, parallel results were obtained for sensitivity and bias. We focus on the sensitivity results below as they are of primary interest.

The full logistic regression model is shown in Table 2.1[4]. Sensitivity was evaluated using the interactions between probe type (target versus foil) and the other predictors (stimulus type and response type). Consistent with the generation effect, probe type interacted with stimulus type such that participants were more likely to correctly recognise studied items in the picture-only than in the picture+word condition (β = 2.01, z = 10.86, 95% CIs [1.63, 2.40], $p < 0.001$). Consistent with the production effect, probe type interacted with response type such that participants were more likely to correctly recognise studied items in the aloud than in the silent condition (β = 1.07, z = 6.36, 95% CIs [0.73, 1.42], $p < 0.001$). An additional interaction between probe type, stimulus type, and response type was also observed, such that studied items in the picture-only condition benefited from production more than those in the picture+word condition (β = 0.74, z = 2.21, 95% CIs [0.06, 1.44], $p < 0.05$).

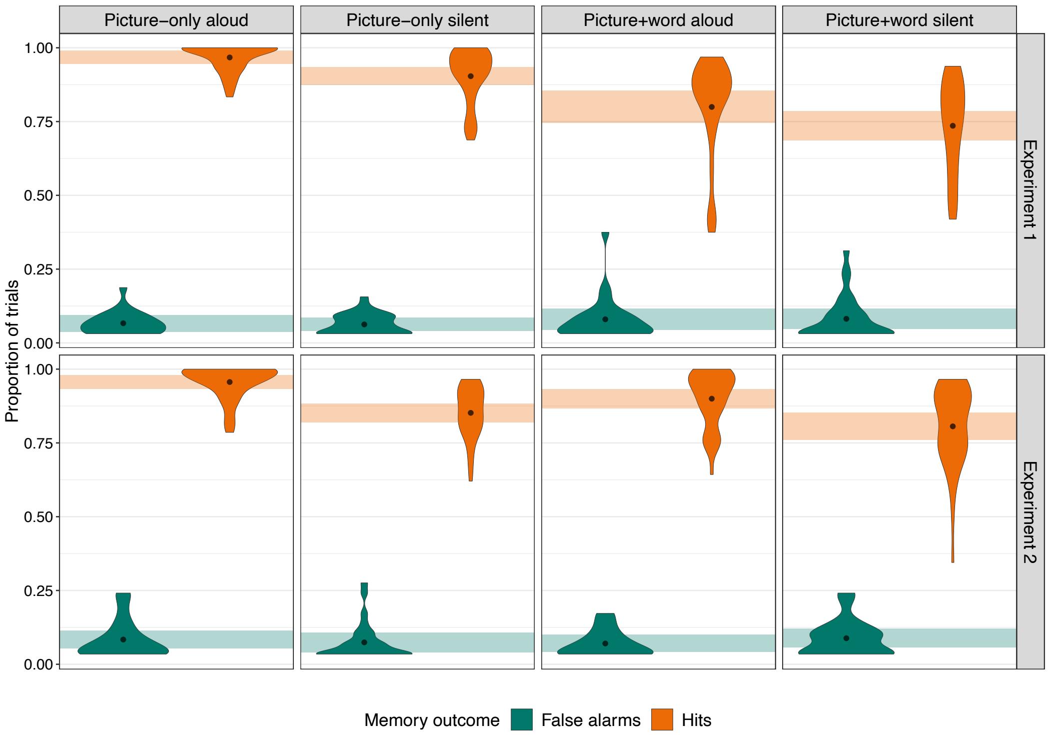


Figure 2.3: Proportions of false alarms and hits for each naming condition (picture-only aloud, picture-only silent, picture+word aloud, picture+word silent) for Experiments 1 (top panel) and 2 (bottom panel). False alarm rates are calculated from the foils in the memory task and hits from the targets. Rectangles represent normalised within-participant 95% confidence intervals.

[4]An additional analysis calculated d' and C. Results were equivalent to the logistic regression and can be seen in Appendix B.

Table 2.1: Experiment 1. Mixed-effects logistic regression of the log-odds of "Yes" responses for all 256 trials. Stimulus type was picture-only or picture+word, Response type was aloud or silent, and Probe type was target or foil.

Fixed effects						
		Estimate	SE	Wald z	p	*CI*
Bias	Intercept	-0.52	0.11	-4.79	< 0.001	-0.75, -0.31
	Stimulus type	0.84	0.11	7.57	< 0.001	0.62, 1.09
	Response type	0.31	0.08	3.62	< 0.001	0.13, 0.48
	Stimulus:Response	0.45	0.17	2.69	< 0.01	0.11, 0.80
Sensitivity	Probe type	5.84	0.22	27.14	0.52	5.42, 6.30
	Probe:Stimulus	2.01	0.18	10.86	< 0.001	1.63, 2.40
	Probe:Response	1.07	0.17	6.36	< 0.001	0.73, 1.42
	Probe:Stimulus:Response	0.74	0.34	2.21	0<.05	0.06, 1.44

Random effects			
		Variance	SD
Participant	Intercept	0.32	0.56
	Probe	1.39	1.18
	Stimulus	0.13	0.36
Item	Intercept	0.40	0.63

Discussion

This experiment yielded two main findings. First, response type impacted memory performance: Saying the picture names aloud during study facilitated picture recognition compared to naming in inner speech. This suggests that better recognition memory stemmed from overt articulation during study. This pattern is consistent with numerous findings in the literature demonstrating a production effect in recognition memory performance (Bodner, Jamieson, Cormack, McDonald, & Bernstein, 2016; Forrin et al., 2012; MacLeod et al., 2010; Mama & Icht, 2016).

Second, participants' recognition performance was impacted by stimulus type, such that it was better when participants had to generate the object names than when the object names were provided as identity primes in the form of written labels superimposed on the pictures. This result may be seen as a generation effect for picture names, similar to generation seen in other paradigms, e.g. when participants read or generate antonyms (Slamecka & Graf, 1978). Naming pictures may be well-practiced, but it is not instantaneous, as it requires conceptual and linguistic processing (Indefrey & Levelt, 2004; Levelt et al., 1999). Although conceptual processing is not necessary to elicit a generation effect, it has been repeatedly shown to be sufficient to do so (Nairne & Widner, 1987; Weldon & Roediger, 1987). This finding is therefore consistent with the literature at large.

However, the interpretation of this finding is complicated by a property of the experimental design. In Experiment 1, all labels correctly indicated the object names. Thus, instead of using the labels to facilitate naming of the pictures, as we assumed, the participants may have entirely relied on the labels and ignored the pictures. This would mean that, despite pictures being present on all trials, the experiment may not have assessed the effect of aided (by identity prime) versus unaided picture naming, but instead may have assessed the effect of word reading versus picture naming. This would entail that the pictures may have been processed less attentively in the picture+word than in the picture-only condition. Therefore, it is possible that this difference in picture processing and not the act of generating the object names can account for the better memory that was found in the picture-only than in the picture+word condition. In other words, the picture superiority effect and not the generation effect may be responsible for the pattern in our data. Experiment 2 was conducted to assess this concern.

We also observed an unanticipated experimental finding: Stimulus type and response type had an overadditive effect such that overt production was more beneficial in the picture-only than in the picture+word condition. This effect is reminiscent of the pattern reported by Fawcett et al. (2012), who found an overadditive effect of stimulus type and response type that they attributed to an interaction between the picture superiority effect and the production effect. If the picture superiority effect did in fact affect our results, then this overadditive effect might also be explained by an interaction between the picture superiority effect and the production effect.

Experiment 2

The goal of Experiment 2 was to minimise the impact of the picture superiority effect by making the picture labels occasionally unreliable. To encourage full processing of the images even in the picture+word conditions, we reassigned 10% of the trials in the naming task to be catch trials. On these trials there was a mismatch between the picture and the label, meaning that participants could no longer be certain that the word they read was a suitable name for the picture they saw. This manipulation was designed to force participants to always look at the pictures in order to evaluate the reliability of the labels with which they were presented, thus making picture processing necessary in all conditions. As a result of this longer visual processing, we expected sensitivity in the picture+word conditions to increase relative to Experiment 1. However, we still predicted effects on sensitivity based upon stimulus type, with higher sensitivity in the picture-only than in the picture+word condition, as well as effects based upon response type, with higher sensitivity in the aloud condition than in the silent condition. Furthermore, if the interaction reported in Experiment 1 was driven by the picture superiority effect, and if the picture superiority effect was now eliminated, we should now see additive, rather than overadditive effects of stimulus type and response type.

Method

Participants

Forty-two participants (33 female, age range: 18-30) who had not participated in Experiment 1 were recruited from the Max Planck Institute participant database and received 10 € for their participation. One participant was excluded because of low performance in the memory test (accuracy <25%), leaving 41 participants in the analysis.

Materials and Design

The same materials and design were used as in Experiment 1 with the exception that 12 items in each picture set were reassigned to be catch trials. This means that there were 116 critical trials and 12 catch trials in the naming task. The labels presented on catch trials were semantically and phonologically unrelated to the pictures. Target picture names and distractor labels were of comparable length (M_T = 7.5, M_D = 6.9), $\log_{10}$ word frequency (M_T = 2.14, M_D = 2.27), familiarity (M_T = 4.27, M_D = 4.35), and manipulability (M_T = 2.8, M_D = 3.0). Three catch items appeared per naming condition.

Catch trials were not included in the memory task. This means that 12 items from each picture set were excluded from the memory test, leaving 232 trials. To ensure that this length difference would not affect our results, data from the first 232 trials in Experiment 1 were reanalysed. Results from the full and shortened datasets did not differ.

Procedure

The procedure was identical to Experiment 1, but for the inclusion of catch trials in the practice and experimental trials.

Analysis

The analysis was identical to Experiment 1.

Naming. As before, naming latencies were calculated to ensure that participants prepared to name the pictures as soon as they were shown. These were compared to the naming latencies of the norming experiment presented in Experiment 1, where participants were instructed to name the pictures as soon as they saw them. In the picture-only condition, norming participants started speaking approximately 1000 ms after picture onset and main study participants started speaking 499 ms (*SD* 188 ms) after the frame (the response cue) was presented. In the picture+word condition, norming participants started speaking approximately 705 ms after picture onset and main study participants started speaking 454 ms (*SD* 137 ms) after the frame was presented. This implies that, as instructed, participants began planning the object names before the response cue was given.

Trials in which participants gave an incorrect response or did not follow the naming instruction were excluded (1.3% of all data). Trials in which participants used the same word to name two objects were excluded (0.2% of all trials), as were trials in which pictures were named incorrectly with the name of a foil (0.1% of all trials).

Memory. A logistic mixed effects model was run using the lme4 package (Bates et al., 2015) in R (version 3.4.0 R Core Team, 2017) with the optimiser *BOBYQA* (Powell, 2009). The same contrasts were used as in Experiment 1. As in Experiment 1, the maximal model failed to converge so the interaction terms for the by-subject and by-item random slopes were removed. Although this model converged, the random slopes for stimulus type and response type effects on participants were highly correlated with the random intercept for participants (-0.96 and 0.99 respectively), indicating overfitting, and were therefore both excluded from the final model. This model included random intercepts for participants and items and random slopes for the effects of stimulus and response type on items, plus a random slope for the effect of probe type on participants. As in Experiment

1, p-values were obtained using likelihood ratio tests and 95% confidence intervals were obtained using the *profile* method of the *confint* function.

Results

Accuracy in the memory task was overall high (91%). Hit rates were higher in the picture-only and the aloud conditions than in the picture+word and the silent conditions (Figure 2.3). As before, the false alarm rates were unaffected by condition, meaning that the bias and sensitivity measures follow identical patterns. For brevity, we again discuss sensitivity only.

The full logistic regression model is shown in Table 2.2. Sensitivity was measured using the interactions between probe type and the other two predictors (stimulus type and response type). As in Experiment 1 there was an interaction between probe type and stimulus type ($\beta = 0.64$, $z = 3.10$, 95% CIs [0.24, 1.05], = $p < 0.001$), meaning that participants were significantly more likely to correctly recognise items in the picture-only than in the picture+word condition. There was also an interaction between probe type and response type ($\beta = 1.49$, $z = 7.13$, 95% CIs [1.09, 1.91], $p < 0.001$), meaning that participants were more likely to correctly recognise items in the aloud than in the silent condition. The interaction between probe type, stimulus type, and response type was not significant ($\beta = -0.09$, $z = -0.24$, 95% CIs [-0.81, 0.64], $p > 0.9$).

Discussion

Experiment 2 differed from Experiment 1 only in the inclusion of catch trials with incongruent labels in the study phase. The purpose of the catch trials was to make it impossible for participants to exclusively rely on the labels for naming and thus force them to process the pictures in both the picture-only and the picture+word condition. This should make the extent to which generation takes place the main difference between stimulus types. The results largely replicate those of Experiment 1. We observed an effect of stimulus type on sensitivity, with better recognition memory when the name of the picture needed to be generated without rather than with the support of a matching label. This supports the existence of a generation effect in picture naming. We also observed an effect of response type on sensitivity, with better recognition memory in the aloud than in the silent condition, consistent with the existence of a production effect

Table 2.2: Experiment 2. Mixed-effects logistic regression of the log-odds of "Yes" responses for all 232 trials. Stimulus type was picture-only or picture+word, Response type was aloud or silent, and Probe type was target or foil.

Fixed effects						
		Estimate	SE	Wald z	p	*CI*
Bias	Intercept	−0.47	0.12	−4.00	< 0.001	−0.72, −0.24
	Stimulus type	0.34	0.09	3.77	< 0.001	0.16, 0.52
	Response type	0.49	0.09	5.45	< 0.001	0.31, 0.66
	Stimulus:Response	0.59	0.17	3.37	< 0.001	0.25, 0.93
Sensitivity	Probe type	5.86	0.20	28.85	< 0.001	5.47, 6.28
	Probe:Stimulus	0.64	0.21	3.10	< 0.01	0.24, 1.05
	Probe:Response	1.50	0.21	7.13	< 0.001	1.09, 1.91
	Probe:Stimulus:Response	−0.09	0.37	−0.24	0.81	−0.82, 0.64

Random effects			
		Variance	SD
Participant	Intercept	0.39	0.62
	Probe	1.03	1.01
Item	Intercept	0.38	0.61
	Stimulus	0.14	0.38
	Response	0.09	0.30

in picture naming. These effects were additive: the interaction reported in Experiment 1 was not replicated.

General discussion

The goal of the present study was to test the hypothesis that producing picture names aloud enhances memory not only by the act of speaking but also by the process of actively generating labels. That is, the generation of picture names improves memory above and beyond the benefit coming from the distinctiveness of the visual features of pictures. In two experiments, we manipulated stimulus type and response type, such that participants generated or read picture names aloud or silently. Our design had two primary methodological differences from Fawcett et al. (2012). First, we used a picture+word condition instead of a word condition. This change aimed to neutralise the visual distinctiveness of pictures and assess whether the generation of labels still improved memory. Second, we used delayed naming to ensure that participants were engaging fully in the task of generating names, even when they were not expected to make an overt response. In both experiments, we showed that both stimulus type (picture-only vs. picture+word) and response type (aloud vs. silent) affected memory performance. This was evidenced by higher sensitivity scores for the picture-only compared to the picture+word conditions, consistent with the generation effect, and higher sensitivity scores for the aloud than for silent conditions, consistent with the production effect.

Experiment 1 also revealed an overadditive effect of stimulus type and response type on sensitivity such that responding aloud benefited items more in the picture-only than in the picture+word condition. Our aim in Experiment 1 was to test whether a generation effect could be captured in picture naming, and therefore we included pictures in all conditions. However, it is possible that the reliability of the labels in the picture+word condition led attention to be directed away from the pictures, making the visual features of the picture more prominent in the picture-only than in the picture+word condition and serving to enhance memory. We addressed this concern in Experiment 2 with the inclusion of unreliable labels, i.e. labels other than the names of the pictures they accompanied. The fact that participants could no longer exclusively rely on the labels meant that they always had to attend to the pictures. We expected this to reinforce visual processing of the pictures and thus improve performance in the picture+word condition. Hit rates in the picture+word condition were higher in Experiment 2 than in Experiment 1, supporting this claim. The interaction between stimulus type and response type disappeared in Experiment 2, consistent with our characterization of this interaction as resulting from differential picture processing in the picture+word condition across experiments.

To ensure the validity of the comparison of sensitivity across Experiments 1 and 2, we ran a logistic regression with experiment as a fixed effect on the pooled data from both experiments (N = 82). Experiment 1 was coded as -0.5 and Experiment 2 as 0.5; the other effects were coded as previously. Results of this analysis are shown in Table 2.3. Importantly, the interaction between probe type and stimulus type, which captures sensitivity towards the two different types of stimuli, changed significantly as a function of experiment. We interpret this as a result of the increase in hit rates in the picture+word condition in Experiment 2, evident in Figure 2.3. Despite this change, sensitivity remained greater for items in the picture-only condition than in the picture+word condition both in Experiment 2 and in the pooled data from Experiments 1 and 2. This provides strong evidence that generation in a picture naming task enhances recognition memory. That is, the process of generating a label for a picture at encoding enhances memory at a later test.

Table 2.3: Comparison of Experiments 1 and 2. Mixed-effects logistic regression of the log-odds of "Yes" responses for first 232 trials. Stimulus type was picture-only or picture+word, Response type was aloud or silent, and Probe type was target or foil.

Fixed effects					
		Estimate	SE	Wald *z*	*p*
Bias	Intercept	−0.50	0.09	−5.82	< 0.001
	Stimulus type	0.58	0.08	7.61	< 0.001
	Response type	0.38	0.07	5.40	< 0.001
	Stimulus:Response	0.52	0.12	4.33	< 0.001
	Experiment	0.07	0.15	0.48	0.64
	Experiment:Stimulus	−0.38	0.13	−2.88	< 0.01
	Experiment:Response	0.22	0.13	1.72	0.09
	Experiment:Stimulus:Response	0.14	0.24	0.56	0.59
Sensitivity	Probe type	5.89	0.15	39.82	< 0.001
	Probe:Stimulus	1.36	0.15	8.90	< 0.001
	Probe:Response	1.41	0.15	9.44	< 0.001
	Probe:Stimulus:Response	0.33	0.26	1.31	0.21
	Probe:Experiment:Stimulus	−1.53	0.28	−5.55	< 0.001
	Probe:Experiment:Response	0.33	0.27	1.23	0.23
	Probe:Experiment:Stimulus:Response	−0.73	0.48	−1.52	0.14

Random effects			
		Variance	SD
Participant	Intercept	0.29	0.54
	Probe	1.20	1.10
	Stimulus	0.06	0.23
	Response	0.04	0.20
	Experiment	0.34	0.58
Item	Intercept	0.39	0.62
	Stimulus	0.17	0.41
	Response	0.05	0.23
	Experiment	0.05	0.23

In contrast to the effects of stimulus type, which varied between experiments, response type showed consistent patterns across both experiments, clearly demonstrating that overt production of picture names greatly improves recognition memory for the pictures compared to production in inner speech. This production effect for picture names was first reported by Fawcett et al. (2012) in three experiments that compared mouthing to inner speech. The two experiments reported here compared overt speech to inner speech and again found a production effect (see also Richler, Palmeri, & Gauthier, 2013). The results of these studies suggest that the production effect for pictures, like the production effect for words, is robust across different means of production.

This study is the first to report a simultaneous role for the production effect and the generation effect in picture naming. This converges with evidence provided by Experiment 1 of Forrin et al. (2014), in which the production and the generation effect independently improved recognition memory for words. It also converges with a larger body of earlier work. For instance, Fawcett et al. (2012) had previously reported a production effect in picture naming by showing that pictures that were named overtly were remembered better than pictures that were named in inner speech. They further showed that pictures are remembered better than words, although it is not possible to discriminate between the picture superiority effect and the generation effect from this comparison. Furthermore, Weldon and Roediger (1987) in their Experiment 2 used a picture-only versus picture+word contrast and found better memory in the picture-only condition, indicative of a generation effect. However, in that experiment, all words were produced in inner speech. The present study combined the methodology of previous work to find that the production effect and the generation effect independently improve memory in picture naming.

The findings of this study have broader implications regarding the generation effect and the way it relates to the picture superiority effect. Regarding the generation effect, we have extended its boundaries to include the naming of intact pictures. Previously, Kinjo and Snodgrass (2000) reported a generation effect for pictures by comparing the naming of fragmented pictures to the naming of intact ones. Here we have shown that the naming of intact pictures, which also requires generation in the form of conceptual or linguistic processing, is able to give similar memory benefits. This corroborates the findings reported by Weldon and Roediger (1987) in their Experiment 2 and underscores the role of picture naming in the generation effect in addition to its role in the picture superiority effect. We have also shown that the generation affect and the picture superiority effect are very closely linked in picture naming tasks, which has implications for the methodology used to study the picture superiority effect. That is, studies primarily interested in the visual distinctiveness of pictures need to carefully control for the differences in the conceptual and linguistic processing required for pictures versus words.

In the Introduction, we noted that the production effect, the picture superiority effect, and the generation effect are often all ascribed to enhanced distinctiveness of the memory representations of the target items. If a common processing principle underlies all effects, one might expect the effects to interact; and indeed this prediction is often borne out (e.g. Fawcett et al., 2012). In Experiment 1 of the present study, an interaction of stimulus type and response type was seen, but this interaction was not replicated in Experiment 2. We interpret this pattern as indicating that the stimulus type effect of Experiment 1 was a compound of the generation effect and the picture superiority effect. When the picture superiority component was eliminated in Experiment 2, the interaction disappeared. Consequently, it appears that in our study, the production effect may have interacted with the picture superiority effect, whereas the production effect and generation effect did not interact. This pattern may be seen to challenge the view that the latter two effects originated from the involvement of a shared mechanism. Alternatively, similar processing mechanisms may be implicated but applied to distinct representations. Thus, the picture superiority effect (not directly assessed in the present study) may be due to an enhancement of the distinctiveness of the visual representation of the target (Nelson et al., 1977); the production effect may arise due to increased distinctiveness of the phonetic representation of the target name and the associated articulatory and motor commands (Forrin et al., 2012); and the generation effect in picture naming may arise from increased distinctiveness of the conceptual representation associated with the picture name. Note that the generation effect in picture naming discussed here for the first time is different from the 'classic' generation effect for words. The participants do not generate the stimulus (the picture) as they do in the classic case, where a probe ("hot") elicits a target ("cold"); instead they generated conceptual and lexical representations that were associated with the targets. Our results show that these processes enhance memory for the target pictures. Exactly how this effect arose needs to be elucidated in further work.

Our study also has implications for psycholinguistic research on the interface of language and memory and, more specifically, for the finding that speakers tend to remember their own utterances better than their listeners do (Hoedemaker et al., 2017; Knutsen & Le Bigot, 2014; McKinley et al., 2017; Yoon et al., 2016). Although in the present study we only tested people individually, we report a similar finding: When participants acted as speakers, i.e. when they generated words or when they spoke aloud, their memory improved relative to when they did not engage in these activities. Thus, the current work provides a useful starting point for research into memory asymmetries between speakers and listeners in conversation.

Conclusion

Generating a word from a cue improves memory for that item relative to reading. In two experiments, we demonstrate that a similar memory benefit arises from a picture naming task. Producing the name of the picture aloud improves memory (the production effect), as does generating a label for the picture (the generation effect). This demonstrates the interplay between language and memory and has implications for both memory and psycholinguistic research.

Appendix A

Below are the two stimuli sets used in Experiments 1 and 2. Word frequency (WF) scores, specifically log_{10} word frequency, comes from the SUBTLEX-NL database. Familiarity (Fam), Visual complexity (VC) and Manipulability (Man) scores come from the BOSS database norms. Naming agreement (Agr.) scores come from our own norming study described in Experiment 1. Starred (*) items were used in the catch trials in Experiment 2.

Set	Filename	Dutch	Agr.	WF	Fam	VC	Man
A	feather03a.jpg	veer	100%	2.18	3.90	2.50	2.20
A	microwave.jpg	magnetron	100%	2.18	4.69	2.19	2.97
A	parrot01.jpg	papegaai	100%	2.16	4.10	3.19	2.18
A	ring01.jpg	ring	100%	3.36	4.30	2.20	3.20
A	iceskate.jpg*	schaats	100%	1.46	4.00	3.00	3.10
A	sewingmachine01a.jpg	naaimachine	100%	1.26	3.90	3.60	2.60
A	hammer01.jpg	hamer	100%	2.57	4.20	2.20	3.80
A	cigar.jpg	sigaar	88%	2.63	4.12	1.70	3.91
A	ant.jpg	mier	100%	2.05	4.62	2.67	1.65
A	pear01.jpg	peer	100%	1.94	4.50	2.10	1.50
A	hen.jpg	kip	100%	3.22	4.43	3.12	2.38
A	shower.jpg	douche	100%	2.99	4.57	2.69	3.56
A	paperclip03.jpg	paperclip	100%	1.46	4.50	2.00	2.00
A	bridge.jpg	brug	100%	3.29	4.52	2.33	2.21
A	surfboard.jpg	surfplank	50%	1.57	3.61	1.85	3.38
A	barrel01.jpg	ton	75%	2.78	4.14	2.07	1.76
A	saw02b.jpg	zaag	100%	2.19	3.70	2.20	4.00
A	corn02.jpg	mais	63%	1.53	4.81	2.40	3.15
A	tshirt.jpg	t-shirt	88%	2.53	4.60	1.70	2.90
A	sheep.jpg	schaap	100%	2.46	4.43	2.90	2.09
A	plate01b.jpg	bord	88%	3.08	4.60	1.70	2.40
A	chessboard.jpg*	schaakspel	63%	1.46	4.20	3.80	2.80
A	celery.jpg	bleekselderij	50%	0.60	4.40	2.20	1.50
A	egg01a.jpg	ei	100%	2.85	4.24	1.93	2.21
A	mushroom01.jpg	champignon	100%	1.15	4.60	2.30	1.60
A	microphone01.jpg	microfoon	100%	2.66	4.60	2.26	4.21
A	helicopter.jpg	helikopter	100%	2.98	4.24	3.45	2.24
A	glass02a.jpg	beker	100%	2.59	4.60	1.80	2.20

A	champagne.jpg	champagne	88%	3.18	4.30	2.70	3.20
A	mirror02.jpg	spiegel	100%	3.08	4.52	2.05	3.32
A	lipstick02a.jpg	lippenstift	100%	2.45	4.10	2.00	4.30
A	boot02b.jpg	laars	100%	2.26	4.40	2.30	2.80
A	hanger02a.jpg	kleerhanger	63%	1.30	4.50	1.80	2.90
A	woodenshoe.jpg	klomp	100%	1.60	3.03	2.07	2.19
A	computermouse06.jpg	muis	100%	2.69	4.80	2.60	3.20
A	flag.jpg	vlag	100%	2.89	4.21	1.83	2.18
A	endive.jpg	witlof	88%	0.78	3.70	2.10	1.60
A	dolphin01.jpg*	dolfijn	100%	1.92	4.48	2.83	1.82
A	hourglass.jpg	zandloper	100%	1.72	3.80	2.80	2.30
A	moon.jpg	maan	88%	3.27	4.37	2.51	2.27
A	panda.jpg	panda	75%	1.63	4.55	3.02	1.85
A	escalator.jpg	roltrap	100%	1.61	4.81	2.64	3.00
A	cork02.jpg*	kurk	100%	1.86	4.33	1.57	3.00
A	lighthouse.jpg	vuurtoren	100%	2.17	4.10	2.71	1.91
A	zebra.jpg	zebra	100%	2.13	4.40	3.05	1.47
A	beerbottle.jpg	bierflesje	75%	1.11	4.64	1.55	3.41
A	tire.jpg	band	50%	3.54	4.62	1.60	2.56
A	cloud.jpg	wolk	100%	2.38	4.74	2.12	2.21
A	wheelchair.jpg	rolstoel	100%	2.56	4.52	2.79	3.85
A	hose.jpg	tuinslang	100%	1.53	4.64	1.93	3.12
A	box01a.jpg*	doos	100%	3.22	3.70	2.50	2.70
A	asparagus.jpg	asperges	100%	1.59	4.00	2.30	1.50
A	lime.jpg	limoen	100%	1.83	3.70	2.20	1.90
A	chimney.jpg	schoorsteen	100%	2.28	4.38	2.07	1.82
A	potato02b.jpg*	aardappel	100%	2.17	4.50	1.90	1.60
A	strawberry.jpg	aardbei	100%	1.84	4.60	2.10	1.50
A	flashlight02b.jpg	zaklamp	100%	2.35	4.30	2.60	3.00
A	microscope.jpg	microscoop	100%	1.93	3.80	3.60	3.50
A	balloon01b.jpg	ballon	100%	2.37	4.40	1.90	3.30
A	pencil01.jpg	potlood	100%	2.38	4.70	1.90	3.50
A	bowl01.jpg	schaal	75%	2.63	4.40	2.10	1.80
A	cane.jpg	wandelstok	88%	1.54	3.50	2.30	3.70
A	stapler03a.jpg	nietmachine	100%	1.53	4.50	2.30	3.50
A	ginger.jpg	gember	88%	1.53	3.80	2.20	1.40
A	spoon01.jpg	lepel	88%	2.34	4.60	2.30	3.40

A	saxophone.jpg	saxofoon	100%	1.64	4.19	3.21	4.18
A	ducttape.jpg*	ductape	75%	0.30	4.55	1.79	2.97
A	broccoli01a.jpg	broccoli	100%	2.07	4.70	2.70	1.30
A	boxcutter03a.jpg	stanleymes	88%	0.90	4.20	2.90	2.80
A	onion.jpg	ui	100%	2.01	4.60	2.20	2.00
A	axe01.jpg	bijl	88%	2.61	3.70	2.40	4.00
A	table01.jpg	tafel	100%	3.56	4.79	1.71	2.88
A	funnel.jpg	trechter	88%	1.38	4.63	1.59	2.71
A	sieve01b.jpg	zeef	75%	1.42	4.20	2.30	2.20
A	watermelon01.jpg	watermeloen	75%	1.88	4.60	2.20	2.20
A	steeringwheel.jpg	stuur	100%	3.70	4.76	2.15	4.32
A	kite.jpg*	vlieger	100%	2.22	4.43	2.10	3.71
A	zucchini01.jpg	courgette	100%	1.20	4.40	2.00	1.50
A	paintbrush01.jpg	kwast	100%	1.88	4.10	2.50	3.50
A	carrot01.jpg	wortel	100%	2.43	4.40	1.90	2.10
A	crown.jpg	kroon	100%	2.80	4.43	2.74	3.44
A	kangaroo.jpg	kangoeroe	100%	1.84	4.21	2.90	2.53
A	magnifyingglass01b.jpg	vergrootglas	63%	1.53	3.40	2.20	3.60
A	handcuffs.jpg*	handboeien	88%	2.66	4.48	2.24	3.91
A	tupperware03a.jpg	bakje	88%	1.81	4.40	2.20	2.00
A	seaturtle.jpg	schildpad	100%	2.28	4.40	3.36	2.15
A	drumset.jpg	drumstel	100%	1.83	4.71	2.95	4.29
A	tree.jpg	boom	100%	3.36	4.69	2.93	2.62
A	banana01.jpg	banaan	100%	2.37	4.70	2.00	3.40
A	frenchfries.jpg	friet	88%	2.17	4.93	1.95	2.71
A	gift01.jpg	cadeau	100%	3.11	4.71	2.40	3.15
A	humanskeleton.jpg	skelet	88%	2.10	4.71	3.31	2.26
A	bread.jpg	brood	100%	3.17	4.34	2.39	2.03
A	screwdriver04b.jpg	schroevendraaier	100%	1.99	4.10	2.30	3.50
A	calculator01.jpg	rekenmachine	100%	1.42	4.30	3.10	3.00
A	elbow.jpg	elleboog	88%	2.16	4.90	1.88	3.44
A	candle08b.jpg	kaars	100%	2.41	4.30	2.20	2.70
A	ladder.jpg	ladder	100%	2.63	4.60	1.93	4.12
A	button01.jpg	knoop	100%	2.62	4.40	1.60	2.60
A	umbrella04.jpg	paraplu	100%	2.18	4.50	2.50	4.20
A	thimble.jpg	vingerhoedje	88%	1.04	3.30	2.30	2.50
A	ruler04.jpg	liniaal	100%	1.40	4.40	2.10	3.10

A	pen04b.jpg	pen	100%	2.98	4.80	2.10	4.00
A	kiwi03.jpg	kiwi	100%	1.46	4.71	2.38	1.91
A	radiator.jpg	verwarming	75%	2.34	4.38	2.43	1.82
A	spatula03.jpg	spatel	100%	1.32	4.40	2.10	2.50
A	grater01a.jpg	rasp	75%	0.85	4.30	2.70	3.30
A	leek.jpg	prei	100%	1.11	3.80	2.50	1.50
A	ironingboard01.jpg	strijkplank	100%	0.95	4.52	1.57	3.62
A	horse.jpg	paard	100%	3.56	4.45	2.88	3.15
A	lettuce.jpg	sla	100%	3.56	4.60	3.00	1.30
A	scale01a.jpg	weegschaal	100%	1.91	4.10	2.90	2.40
A	bed.jpg	bed	100%	4.02	4.80	1.98	3.74
A	shoppingcart.jpg	winkelwagen	75%	1.20	4.71	2.38	3.32
A	tulip02.jpg	tulp	100%	1.34	4.40	2.12	2.35
A	pumpkin.jpg	pompoen	100%	2.04	4.71	2.29	2.18
A	glasses01a.jpg*	bril	100%	3.03	4.30	2.80	3.90
A	ear.jpg	oor	100%	3.04	4.95	2.43	3.88
A	coconut.jpg	kokosnoot	88%	1.88	3.90	2.30	1.70
A	rollingpin01a.jpg	deegroller	100%	0.60	3.80	2.10	3.60
A	pizza.jpg*	pizza	100%	3.03	4.40	2.50	2.70
A	sock01a.jpg	sok	100%	2.14	4.50	2.00	3.40
A	hairdryer02a.jpg	fohn	100%	0.60	4.20	2.90	4.00
A	shoulder.jpg*	schouder	88%	2.91	4.93	2.48	3.29
A	anchor.jpg	anker	100%	2.34	4.20	1.85	2.32
A	rockingchair.jpg	stoel	63%	3.35	4.69	2.40	3.68
A	violin.jpg	viool	100%	2.28	3.60	3.20	4.40
A	smokingpipe.jpg	pijp	100%	2.78	4.10	1.79	3.65
B	pillow01a.jpg	kussen	100%	3.30	4.40	2.40	3.50
B	nailclipper03b.jpg	nagelknipper	75%	0.85	4.40	2.60	3.90
B	shoelace.jpg	veter	100%	1.77	4.30	1.90	3.10
B	battery02b.jpg	batterij	100%	2.47	4.60	2.20	1.60
B	raspberry01.jpg	framboos	100%	1.11	4.70	2.30	1.60
B	suitcase.jpg*	koffer	100%	3.17	4.20	2.70	2.80
B	teabag.jpg	theezakje	100%	0.95	4.40	2.40	2.80
B	lighter01.jpg	aansteker	100%	2.40	4.20	2.70	3.80
B	remotecontrol04.jpg	afstandsbediening	88%	2.42	4.40	3.20	3.60
B	belt.jpg	riem	100%	2.79	4.40	2.00	3.80
B	razor01.jpg	scheermesje	100%	1.57	4.30	2.50	3.70

B	helmet.jpg	helm	88%	2.69	3.57	2.02	2.32
B	snowman.jpg*	sneeuwpop	100%	1.68	4.62	2.26	3.26
B	blender.jpg	blender	88%	1.49	4.52	2.29	2.79
B	tie02.jpg	stropdas	100%	2.29	4.64	1.52	4.21
B	cow.jpg	koe	100%	2.91	4.71	2.90	2.91
B	tomato01.jpg	tomaat	100%	2.12	4.70	1.70	1.50
B	bathtub.jpg	bad	75%	3.27	4.79	2.26	3.32
B	avocado01.jpg	avocado	88%	1.38	4.30	2.50	1.50
B	uprightpiano01.jpg	piano	100%	2.79	4.67	2.98	4.47
B	couch02.jpg	bank	100%	3.60	4.81	2.12	3.21
B	peanut01.jpg	pinda	100%	1.97	4.20	2.10	1.90
B	towel01.jpg*	theedoek	50%	1.23	4.30	2.10	2.20
B	safetypin.jpg	veiligheidsspeld	63%	1.04	4.20	1.90	2.20
B	fork03c.jpg	vork	100%	2.36	4.60	2.20	3.50
B	road02.jpg	weg	100%	4.81	4.79	1.69	2.59
B	pencilsharpener02a.jpg	puntenslijper	100%	0.30	4.30	2.50	3.40
B	pepper04a.jpg	paprika	100%	1.77	4.60	2.20	1.20
B	computerkeyboard02.jpg	toetsenbord	100%	1.69	4.70	3.10	4.00
B	clothespin03b.jpg	wasknijper	75%	0.30	4.30	2.30	2.60
B	match.jpg	lucifer	100%	2.45	4.40	1.90	3.80
B	treadmill.jpg	loopband	75%	1.20	4.67	2.62	3.59
B	cigarette.jpg	sigaret	100%	3.09	4.10	2.00	4.20
B	apple07.jpg	appel	100%	2.65	4.60	1.80	2.40
B	strainer02.jpg	vergiet	100%	1.59	4.20	2.10	2.00
B	africanelephant.jpg	olifant	100%	2.72	4.48	2.95	2.68
B	aluminiumfoil.jpg	aluminiumfolie	86%	1.42	4.79	1.88	2.65
B	fingerprint.jpg	vingerafdruk	100%	2.28	4.69	3.17	3.26
B	whisk.jpg	garde	100%	2.32	3.80	2.50	2.80
B	giraffe.jpg	giraffe	100%	1.75	4.43	3.12	2.21
B	bicycle.jpg	fiets	88%	2.98	4.83	2.88	4.18
B	vase01.jpg	vaas	100%	2.30	3.70	2.50	2.10
B	stool01.jpg	kruk	100%	2.05	4.67	1.81	3.12
B	mattress.jpg*	matras	100%	2.35	4.86	2.05	3.26
B	zipper.jpg	rits	88%	2.28	4.83	2.14	4.24
B	vacuumcleaner01.jpg	stofzuiger	100%	2.03	4.50	2.55	3.79
B	puzzlepiece.jpg	puzzelstukje	100%	1.15	4.30	2.70	1.90
B	television.jpg*	televisie	88%	2.99	4.76	1.78	3.56

B	car.jpg	auto	100%	4.30	4.57	2.98	3.32
B	iron01b.jpg	strijkijzer	88%	1.51	4.40	3.00	3.90
B	broom01.jpg	bezem	88%	2.22	4.30	2.30	3.90
B	bikepump01.jpg	fietspomp	86%	0.70	4.52	1.88	3.65
B	rug01.jpg	kleed	75%	3.02	4.10	2.50	2.00
B	thermometer02b.jpg	thermometer	88%	1.76	4.10	2.70	1.80
B	tweezers02a.jpg	pincet	100%	1.60	4.30	1.90	3.30
B	sandcastle.jpg	zandkasteel	100%	1.11	4.36	2.60	2.09
B	rice.jpg*	rijst	100%	2.60	4.50	2.40	1.50
B	scissors01.jpg	schaar	100%	2.45	4.50	2.50	4.20
B	safe.jpg	kluis	100%	3.19	4.26	2.36	3.38
B	saltshaker03a.jpg	zout	50%	2.83	4.76	1.74	3.59
B	fan.jpg	ventilator	75%	2.05	4.20	3.10	2.40
B	bandaid01.jpg*	pleister	88%	2.03	3.90	1.90	3.20
B	straw.jpg*	rietje	100%	2.06	3.70	1.80	3.70
B	dishsoap.jpg	afwasmiddel	100%	1.04	4.40	2.30	2.20
B	photocopier.jpg	printer	88%	1.60	4.52	3.10	2.94
B	lemon02.jpg	citroen	88%	2.36	4.60	2.20	2.10
B	pill.jpg	pil	100%	2.61	4.00	1.90	2.70
B	pacifier02d.jpg	speen	88%	1.23	3.80	2.70	3.30
B	pig.jpg*	varken	100%	3.03	4.36	2.57	2.24
B	windmill.jpg	molen	100%	2.30	4.33	2.81	2.00
B	rock01a.jpg	steen	75%	3.19	4.10	2.30	2.30
B	garbagecan02.jpg	prullenbak	75%	1.79	4.57	1.93	3.12
B	accordion01.jpg	accordeon	100%	1.72	4.26	3.33	4.12
B	eggplant.jpg	aubergine	100%	1.49	4.00	2.20	1.50
B	yarn.jpg	wol	75%	2.22	3.70	2.70	2.10
B	leaf02a.jpg*	blad	100%	2.69	4.40	1.80	2.20
B	almond.jpg	amandelen	86%	1.86	4.30	2.10	1.30
B	bow.jpg	boog	100%	2.57	4.20	1.76	4.29
B	bucket01a.jpg	emmer	100%	2.47	3.90	2.00	2.50
B	laptop01a.jpg*	laptop	100%	2.41	4.60	3.60	3.20
B	necklace.jpg	ketting	100%	2.92	4.20	2.50	3.30
B	medal02b.jpg	medaille	100%	2.65	3.80	2.10	2.40
B	envelope03a.jpg	envelop	88%	2.44	4.60	1.80	3.20
B	tray.jpg	dienblad	100%	1.74	3.40	2.30	2.50
B	tiger02.jpg	tijger	100%	2.71	4.36	3.12	1.82

B	greatwhiteshark.jpg	haai	100%	2.62	4.33	2.86	2.12
B	toaster01.jpg	broodrooster	100%	2.03	4.50	2.90	2.70
B	swing.jpg	schommel	100%	1.85	4.61	1.59	4.00
B	cat.jpg	kat	88%	3.36	4.48	2.86	2.79
B	donut.jpg	donut	100%	2.29	4.74	1.71	2.56
B	chalk.jpg	krijt	88%	2.71	3.70	1.70	2.60
B	bus.jpg	bus	100%	3.45	4.69	2.45	2.29
B	cannon.jpg	kanon	100%	2.43	4.26	2.52	2.65
B	key01.jpg	sleutel	100%	3.55	4.88	1.95	3.97
B	coatrack.jpg	kapstok	100%	1.51	4.55	1.69	3.35
B	basketball01.jpg	basketbal	100%	2.59	4.79	1.83	4.26
B	fence02.jpg	hek	88%	3.00	4.38	2.24	1.94
B	watch02a.jpg	horloge	100%	3.09	4.40	3.10	3.80
B	garlic01a.jpg	knoflook	100%	2.29	4.60	2.40	1.60
B	ladybug03.jpg	lieveheersbeestje	100%	1.28	4.69	2.79	1.53
B	englishcucumber.jpg	komkommer	88%	1.77	4.74	1.86	2.06
B	bowlingpin.jpg	kegel	88%	1.30	4.46	1.61	3.50
B	cauliflower01.jpg	bloemkool	100%	1.40	4.40	2.50	1.20
B	jar03.jpg	pot	88%	3.13	3.90	2.10	1.80
B	dice05a.jpg	dobbelsteen	100%	1.51	4.50	2.30	3.40
B	woodboard.jpg	plank	100%	2.69	3.80	2.20	1.50
B	shovel01.jpg	schep	63%	2.31	4.71	1.52	3.97
B	tent.jpg	tent	100%	3.25	4.45	2.14	2.82
B	arrow02.jpg	pijl	100%	2.49	4.24	1.71	3.53
B	binoculars01b.jpg*	verrekijker	100%	2.09	4.00	3.50	4.40
B	blackbear.jpg	beer	100%	3.05	4.33	3.07	2.09
B	musicsheet.jpg	bladmuziek	100%	1.11	4.36	2.74	2.94
B	greenolive.jpg	olijf	88%	1.53	4.30	2.00	1.50
B	weight01.jpg	gewicht	75%	2.88	3.80	2.00	3.20
B	bib.jpg	slabbetje	100%	1.23	3.80	2.50	2.70
B	acousticguitar02.jpg	gitaar	100%	2.70	4.52	2.45	4.71
B	pineapple01a.jpg	ananas	100%	2.05	4.50	3.10	1.70
B	book01b.jpg	boek	100%	3.82	4.50	2.20	3.90
B	coffeebean.jpg	koffiebonen	100%	1.20	4.30	2.10	1.90
B	bracelet01.jpg	armband	88%	2.44	3.40	2.20	2.90
B	cactus.jpg	cactus	100%	1.90	4.14	2.74	1.85
B	headphones02b.jpg	koptelefoon	100%	1.96	4.40	2.90	3.60

B	monarchbutterfly.jpg	vlinder	100%	2.43	4.62	3.24	1.91
B	cupcake.jpg	cupcake	75%	1.08	4.81	2.26	2.41
B	knee.jpg	knie	100%	2.65	4.93	2.02	3.82
B	mug05.jpg	mok	75%	1.70	4.88	1.38	3.82
B	eraser.jpg	gum	100%	1.34	4.30	1.90	3.30
B	trampoline.jpg	trampoline	100%	1.53	4.40	1.98	4.06

Appendix B

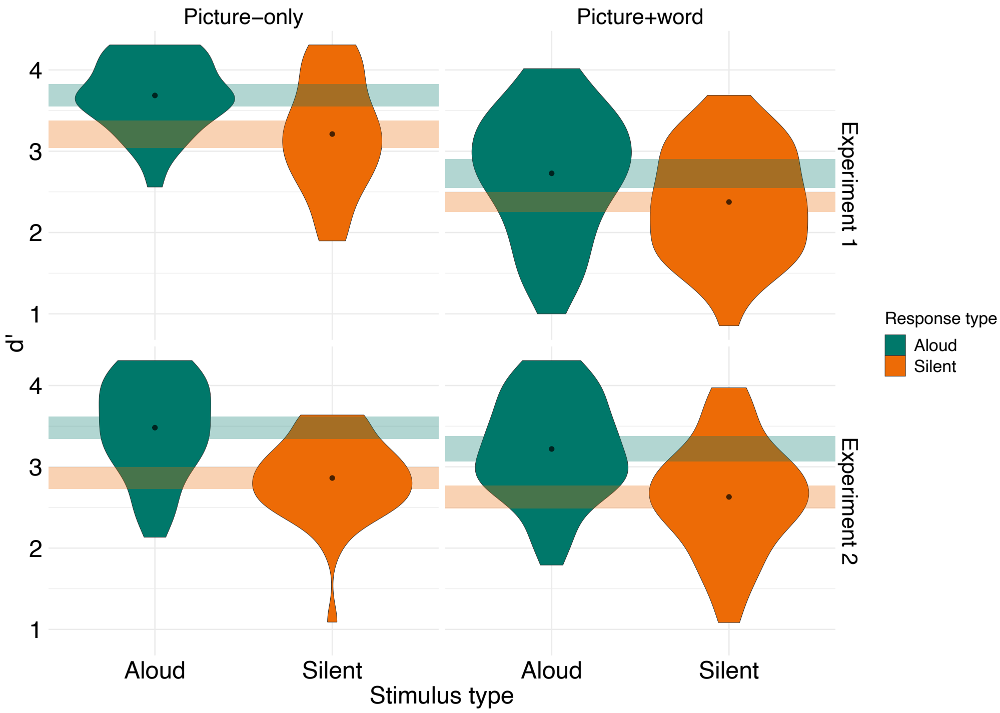


Figure 2.B1: Sensitivity (*d′*) results for each naming condition in Experiments 1 and 2. Columns represent stimulus type; rows represent experiment. Dots represent means by condition. Rectangles represent normalised within-subject 95% confidence intervals by condition.

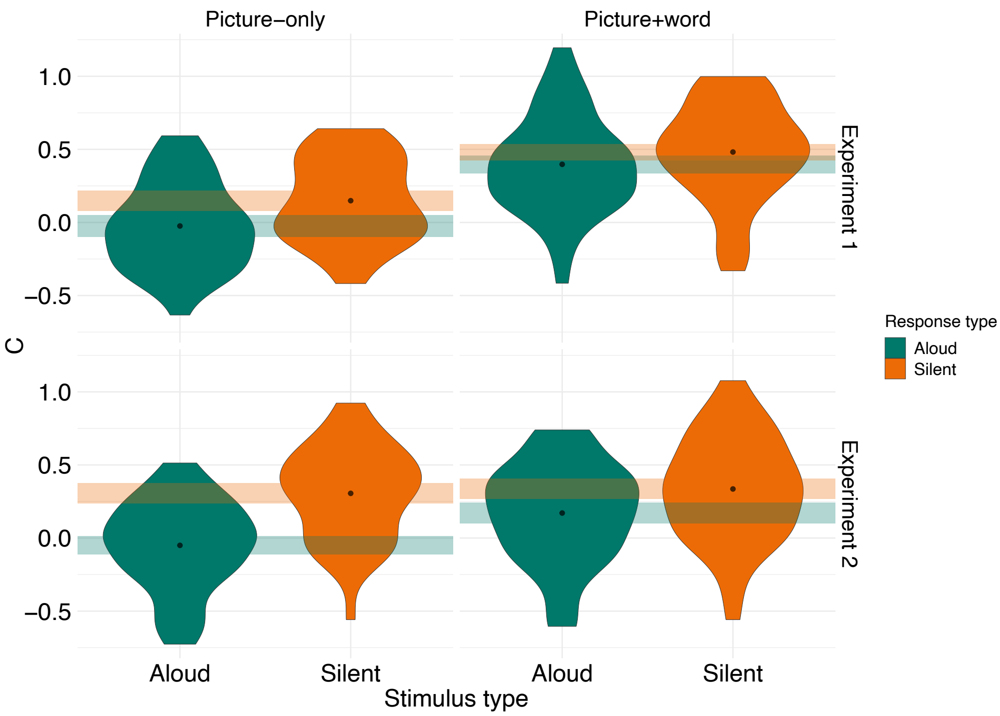


Figure 2.B2: Response bias (*C*) results for each naming condition in Experiments 1 and 2. Columns represent stimulus type; rows represent experiment. Dots represent means by condition. Rectangles represent normalised within-subject 95% confidence intervals by condition.

3 Slow naming of pictures facilitates memory for their names[1]

Abstract

Speakers remember their own utterances better than those of their interlocutors, suggesting that language production is beneficial to memory. This may be partly explained by a generation effect: the act of generating a word is known to lead to a memory advantage (Slamecka & Graf, 1978). In earlier work, we showed a generation effect for recognition of images (Zormpa, Brehm, Hoedemaker, & Meyer, 2019). Here, we tested whether the recognition of their names would also benefit from name generation. Testing whether picture naming improves memory for words was our primary aim, as it serves to clarify whether the representations affected by generation are visual or conceptual/lexical. A secondary aim was to assess the influence of processing time on memory. Fifty-one participants named pictures in three conditions: after hearing the picture name (identity condition), backward speech, or an unrelated word. A day later, recognition memory was tested in a Yes/No task. Memory in the backward speech and unrelated conditions, which required generation, was superior to memory in the identity condition, which did not require generation. The time taken by participants for naming was a good predictor of memory, such that words that took longer to be retrieved were remembered better. Importantly, that was the case only when generation was required: in the no-generation (identity) condition, processing time was not related to recognition memory performance. This work has shown that generation affects conceptual/lexical representations, making an important contribution to the understanding of the relationship between memory and language.

[1]Adapted from Zormpa, E., Meyer, A. S., & Brehm, L. E. (2019). Slow naming of pictures facilitates memory for their names. *Psychonomic Bulletin & Review, 26*, 1675-1682. doi: 10.3758/s13423-019-01620-x.

Introduction

Memory and language are tightly linked. For instance, we have memory representations of the contents of conversations, concerning the facts and events mentioned, but also, often, of the words and phrases used. Additionally, language influences memory: Participants in pair studies remember their own utterances better than their interlocutors' (e.g. Fischer et al., 2015; Hoedemaker et al., 2017; Yoon et al., 2016). Despite this relationship, memory and language are often researched independently. Our goal is to investigate these domains in tandem, exploring the influence of language production on memory. We do so by examining how picture naming, and the time it requires, influences memory for the picture names.

In an earlier study, we asked participants to name pictures with either the picture names or scrambled letter fragments superimposed and later tested their recognition of the pictures (Zormpa, Brehm, et al., 2019). Pictures with the scrambled letters superimposed were remembered better than ones with the picture names superimposed. This can be interpreted as a *generation effect* (Slamecka & Graf, 1978): When participants generated picture labels themselves, they remembered the pictures better than when the correct names were provided. We interpreted this finding under a distinctiveness account (Hunt & Worthen, 2006). Active generation of the names creates an additional episodic memory trace, which aids picture recognition. This framework has successfully explained a range of phenomena in episodic and visual memory, including the picture superiority effect (Paivio et al., 1968) and the production effect (MacLeod et al., 2010).

Our earlier work showed the influence of language production on memory, but it did not allow us to determine what level of representation benefits from generation. Providing a picture label reduces the need to perform object identification, a time-consuming part of picture naming (Indefrey & Levelt, 2004). Therefore, one might expect visual features to be less distinctive in memory when the picture name is provided than when it is not. Another possibility is that generation affects later conceptual and/or lexical processes. For instance, as more time is spent looking at a picture during object recognition and name retrieval, more conceptual and lexical information may become activated. In other words, episodic conceptual and lexical representations may be encoded more strongly when picture labels are generated than when they are provided, leading to more distinctive memory representations. Effects may also arise at multiple levels and interact. To begin to distinguish between these possibilities, we conducted a cross-modal version of our earlier experiment: During study, we presented pictures to be named, but during test, we presented the picture names to be recognised. We reasoned that, as the pictures were not shown again, any generation effect could not be due to better memory for the visual properties of the stimuli. Format changes between training and testing have been shown to not affect memory phenomena like the picture superiority effect (Borges et al., 1977) and the production effect (Mama & Icht, 2016), but the generation effect reported in Zormpa, Brehm, et al. (2019) was only established when both training and test stimuli were pictures.

We used a study-test paradigm involving three study conditions and a Yes/No recognition memory task 24 hours later. The study phase was a self-paced picture naming task where participants heard primes immediately preceding the targets. There were three prime types: the *identity* prime was the target word itself, the *backward* prime was backward speech, and the *unrelated* prime was a word semantically and phonologically unrelated to the target. This paradigm differed from Zormpa, Brehm, et al. (2019) in that testing was delayed and that the primes were auditory (not written). Both changes made the task harder, ensuring performance would not be at ceiling. In addition, primes were now presented before the picture, allowing participants to concentrate on processing the pictures when they appeared. Based on earlier work (e.g. Schriefers, Meyer, & Levelt, 1990), we expected slower naming in the unrelated compared to the backward prime condition because of competition between the targets and unrelated words. This increase in processing time[2] should lead to improved recognition memory for picture names in the unrelated compared to the backward condition. Although semantically related distractors would have elicited even slower naming, such distractors could influence memory in unexpected ways, confounding the effects of processing time and semantic relatedness.

We expected unrelated distractors to have a small effect, hindering testing of our hypothesis that an increase in processing time at input would benefit recognition memory. Therefore, we recorded the picture naming latencies during study and predicted that memory would be best for the

[2] In our preregistration we conceptualised this as cognitive effort and collected pupil sizes for an exploratory study reported at: https://osf.io/w39gu/. However, we were only able to measure processing time, so we restrict our discussion to that.

items that were named the slowest. On study trials, a dot was displayed to the right of the picture. To advance to the next trial, participants fixated the dot and then pressed "Enter". We measured when participants turned from the picture to the dot (gaze durations), and when they pressed the button (button-press latencies) as indices of exposure time used in exploratory analyses. The self-paced naming task served to ensure that the latency measures reflected the time needed for each trial, and the surprise memory task served to avoid rehearsal of the pictures during naming.

Method

Participants

Sixty individuals (48 female, mean age 22.62 years; range 18 - 30 years) participated in this experiment. Eight participants were excluded, two due to technical problems during study and six for not completing the test phase. One additional participant was excluded because of substantially lower Hit rates (approx. 25%) than the other participants. This left data from 51 individuals. Participants were recruited from the Max Planck Institute participant database and received 8 €. All were native Dutch speakers with normal or corrected-to-normal vision; none reported speech or language problems. A power analysis using an effect size of 3% from Zormpa, Brehm, et al. (2019) showed 48 participants would provide sufficient power to answer our main research question. Scripts are at https://osf.io/7rq6n/. Ethical approval was given by the Ethics Board of the Social Sciences Faculty of the Radboud University.

Materials and design

Stimuli were comprised of 246 colour pictures selected from the BOSS database (Brodeur et al., 2010, 2014) presented in 250x250 pixel resolution against a light-gray background (RGB 153, 153, 153). Match software (van Casteren & Davis, 2007) was used to split the pictures into two sets (A and B) matched on name agreement (M_A = 0.93, M_B = 0.95), familiarity (M_A = 4.36, M_B = 4.39), visual complexity (M_A = 2.36, M_B = 2.35), manipulability (M_A = 2.87, M_B = 2.88), $\log_{10}$ word frequency (M_A = 2.15, M_B = 2.22), and duration (ms; M_A = 683, M_B = 679). Familiarity, visual complexity, and manipulability scores were extracted from the BOSS database, frequency from the

SUBTLEX-NL database (Keuleers et al., 2010), and name agreement scores were collected from six native Dutch speakers that did not participate in the experiment. Either set A or B was presented at study; the condition in which they appeared was counterbalanced across six lists. All 246 images were presented at test, such that for three lists, set A served as foils and for the other three set B served as foils.

At study, participants heard a label before each picture[3] (duration: 285-1234 ms) recorded by a female native Dutch speaker. In the identity condition, primes were the picture names. In the backward condition, primes were pseudo-randomly selected foils played backwards, created using the "Reverse" command in Praat (Boersma & Weenink, 2018). None of the backward foils sounded like the targets. In the unrelated condition, primes were semantically and phonologically unrelated Dutch words. Semantic relatedness was judged by a native Dutch speaker. Phonological relatedness was determined by Levenshtein distance. There was no more than 33.3% overlap between targets and unrelated primes on either measure. Unrelated primes were matched between lists, each appearing for a set A and a set B item.

Apparatus and procedure

The study phase was a picture naming task conducted at the Max Planck Institute for Psycholinguistics in a soundproof booth with comfortably dim constant lighting. The experiment was controlled using Presentation (version 18.3; Neurobehavioral Systems, Berkeley, CA, USA) and displayed on a 24" monitor (1920x1080 pixel resolution). The right eye was tracked using an EyeLink 1000 Desktop Mount eye-tracker (SR Research Ltd., Osgoode, Canada) sampling at 500 Hz. A head stabiliser was used to minimise head movements and maintain constant distance between participants' eyes and the screen (54 cm from the end of the camera to the distal end of the chinrest pad). The table height was adjusted for each participant. The experiment began with random-order nine-point calibration and validation routines. Trials began with a drift-check, followed by a white fixation cross

[3]In an earlier experiment, with written primes presented 900-1500 ms before the picture, the generation effect disappeared (https://osf.io/5xe8f/). This may be due to the match between study and test in the identity condition: In this condition, participants saw a written word which they later had to recognise, whereas in the other two conditions participants produced a word in the spoken modality and later had to recognise it in the written modality. Alternatively, the long delay between prime and target may have separated them in two episodes, thus enhancing memory.

displayed on the centre left of the screen (coordinates 480,540) for 700 ms. Then participants heard an audio prime (Sennheiser headphones). At the offset of the prime, a picture replaced the fixation cross and a dot appeared at the centre right of the screen (coordinates 1440,540). Both remained on screen until the trial ended. After participants named the picture, they looked at the dot and pressed "Enter" to move to the next trial. If they had not fixated the dot for minimally 50 ms, nothing happened when pressing "Enter"; this routine served to dissociate gaze duration on the picture from total trial time. Before the experimental trials, participants completed 12 practice trials (four per condition) that provided an opportunity for questions; no feedback was given after this point. Trials from the three conditions were intermixed and presented in a unique random order for each participant. The session lasted 20-25 minutes. The test phase was a self-paced Yes/No memory task conducted online using the LimeSurvey (Version 3.14.8) platform. Links and unique tokens were emailed to participants 20 hours after the study phase. Participants had eight hours to complete the task and were sent reminders after four and six hours if needed. Participants saw 246 words (123 targets, 123 foils), one at a time, and were instructed to press “Ja" (Yes) for words used at study and “Nee" (No) for the remaining words. They moved to the next trial by pressing “Volgende" (Next). The session lasted 10-15 minutes and was followed by debriefing.

Analysis

The main dependent variable was memory performance. “Yes”and “No” responses (coded as 1 and 0 respectively) were analysed using mixed-effects logistic regression. This mirrors standard signal detection analyses while accounting for participant and item variability (DeCarlo, 1998). The main predictors were prime condition and naming latency. All analyses with prime condition as a predictor were run on a dataset containing targets only, as foils did not belong in a prime condition. Latency measures were centred and log-transformed (natural log) to resolve convergence issues. The output of the models with centred or logged and centred latencies followed the same pattern. Contrasts for predictors are described below for each analysis. Unless otherwise specified, analyses were preregistered on the Open Science Framework (https://osf.io/sqad9/).

Analyses were run using the lme4 package (version 1.1-18-1; Bates et al., 2015) in R (version 3.5.0; R Core Team, 2018) with the optimiser *BOBYQA*

(Powell, 2009). Initially, the maximal models were fit and then reduced to overcome convergence problems or overfitting (correlations exceeding 0.95). Reported *p*-values were obtained from maximum likelihood tests comparing a full model to one without the effect of interest. Reported 95% confidence intervals were calculated using the *profile* method of the *confint* function.

In addition to the main analyses on memory performance, we conducted preregistered exploratory analyses using the gaze duration data.

Trials were excluded when participants failed to name a picture, when they named it with an unexpected word, or when they repeated a word (9% of the data). This includes naming two pictures with the same word (e.g."doughnut" for both a doughnut and a bagel) and naming a target (e.g. a shark) using the name of a foil in the memory task (e.g. "dolphin").

Results

Participants' memory was generally accurate (M = 79%, SD = 41%).

The first analysis examined the effect of probe type (target vs. foil) on memory performance, as measured by Yes responses in the memory task, to assess overall accuracy and response bias[4]. This analysis was run separately from the prime condition analysis, as the design of this study was not fully crossed, i.e. foils did not appear in a prime condition. Probe type was sum-to-zero contrast-coded (targets = 0.5, foils = -0.5). The random effects structure included by-participant and by-item intercepts and by-participant and by-item random slopes for probe type.

Results and a visualisation appear in Table 3.1 and Figure 3.1. The significant negative intercept reflects a "No" bias. The significant effect of probe type reflects that participants were more likely to say "Yes" to targets than foils, i.e., they were highly accurate in differentiating between old and new items.

We then examined the effect of prime condition on memory performance. Prime condition was Helmert coded and split into two contrasts. The first contrast tested the effect of generation by comparing the identity condition (contrast=-0.5) to the average of the backward and unrelated conditions (contrast for both=0.25), while the second contrast tested the ef-

[4]The preregistered analysis used accuracy, not "Yes" responses. That model had convergence problems when calculating 95% CIs, presumably due to invariance in the dependent variable. This was solved by using "Yes" responses.

Table 3.1: Mixed-effects logistic regression testing the effect of Probe, i.e., targets vs. foils, on memory (log-odds of "Yes" responses).

Fixed effects					
	Estimate	SE	Wald *z*	*p*	*CI*
Intercept	-0.96	0.13	-7.24	< 0.001	-1.23, -0.70
Targets vs. Foils	3.92	0.24	16.10	< 0.001	3.44, 4.42

Random effects			
		Variance	SD
Participant	Intercept	0.26	0.51
	Targets vs. Foils	1.28	1.13
Item	Intercept	0.76	0.87
	Targets vs. Foils	2.43	1.56

fect of processing time (as a result of competition) by comparing the backward (contrast=-0.5) to the unrelated condition (contrast=0.5). The random effects structure included by-participant and by-item intercepts and by-participant and by-item slopes for the generation contrast.

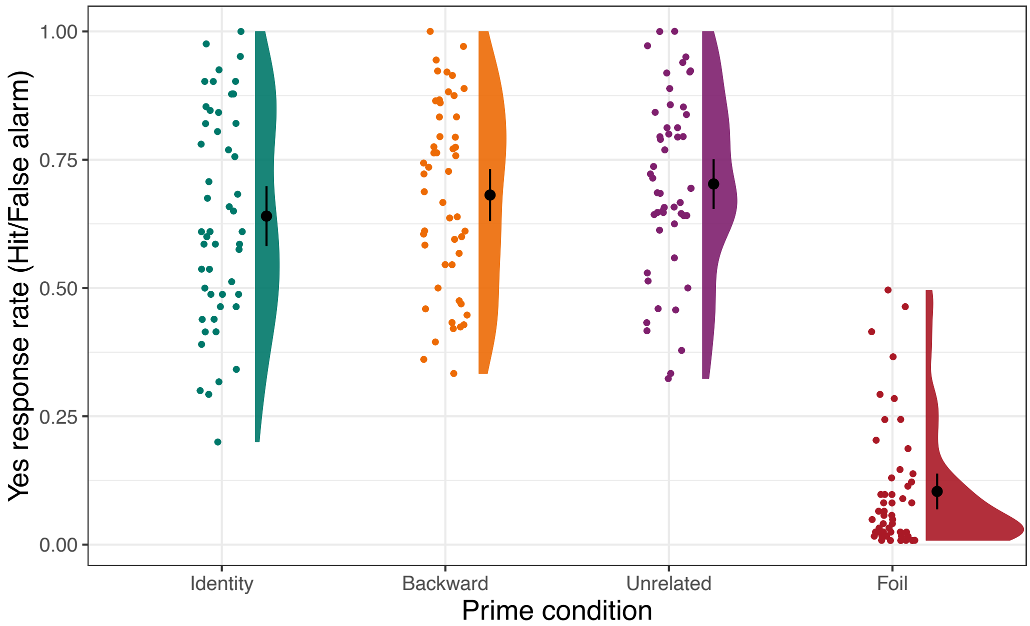
Results appear in Table 3.2 and Figure 3.1. The significant positive intercept term reflects a "Yes" bias to targets, indicating high accuracy. Hit rates were significantly higher in the backward and unrelated conditions than in the identity condition, showing a memory benefit for generated words. In contrast, hit rates did not significantly differ between the backward and unrelated conditions.

Table 3.2: Mixed-effects logistic regression testing the effects of generation and processing time (as manipulated by prime condition) on memory (log-odds of "Yes" responses).

Fixed effects					
	Estimate	SE	Wald *z*	*p*	*CI*
Intercept	1.05	0.16	6.47	< 0.001	0.73, 1.37
Identity vs. Backward & Unrelated	0.46	0.16	2.87	0.01	0.13, 0.78
Backward vs. Unrelated	0.13	0.08	1.53	0.13	-0.04, 0.29

Random effects			
		Variance	SD
Participant	Intercept	1.10	1.05
	Generation	0.69	0.83
Item	Intercept	0.75	0.87
	Generation	0.36	0.60

Figure 3.1: Hit rates by prime condition. The dot represents the condition mean and the bars normalised within-participant 95% confidence intervals.



We then tested how well the three latency measures predicted memory. All measures were positively correlated, with a moderate-to-strong correlation between naming latency and gaze duration ($r = 0.46$, $p < 0.001$), and strong correlations between gaze duration and button-press latency ($r = 0.59$, $p < 0.001$) and between naming and button-press latency ($r = 0.65$, $p < 0.001$).

The prime conditions differed with regard to average naming times: Participants were approximately 300 ms faster in the identity condition than in the backward and unrelated conditions, due to repetition priming (Table 3.3). As such, naming latencies should also be a good predictor of subsequent memory. Additionally, since the three latency measures were highly correlated, the same should hold for the other latency predictors.

Table 3.3: Means and *SDs* (ms) for naming latency, gaze duration, and button-press latency for each prime condition.

	Naming latency	Gaze duration	Button-press latency
Identity	656.43 (172.95)	1138.44 (593.19)	1426.84 (526.46)
Backward	965.93 (355.02)	1423.51 (659.36)	1757.30 (626.35)
Unrelated	983.90 (402.44)	1433.74 (694.92)	1777.36 (658.16)

Table 3.4: Mixed effects logistic regression testing the effect of prime condition and naming latency on memory (log-odds of Yes responses). Naming latency has been log-transformed and centred.

Fixed effects					
	Estimate	SE	Wald z	p	*CI*
Intercept	0.89	0.16	5.61	< 0.001	0.57, 1.20
Naming latency	1.26	0.07	5.74	< 0.001	0.84, 1.72
Identity vs. Backward & Unrelated	0.08	0.13	0.60	0.55	−0.18, 0.33
Backward vs. Unrelated	0.17	0.09	1.92	0.057	−0.01, 0.35
Naming latency:Id. vs. Bw. & Un.	2.30	0.41	5.68	< 0.001	1.51, 3.13
Naming latency:Bw. vs Un.	−0.68	0.33	−2.08	0.04	−1.34, −0.02

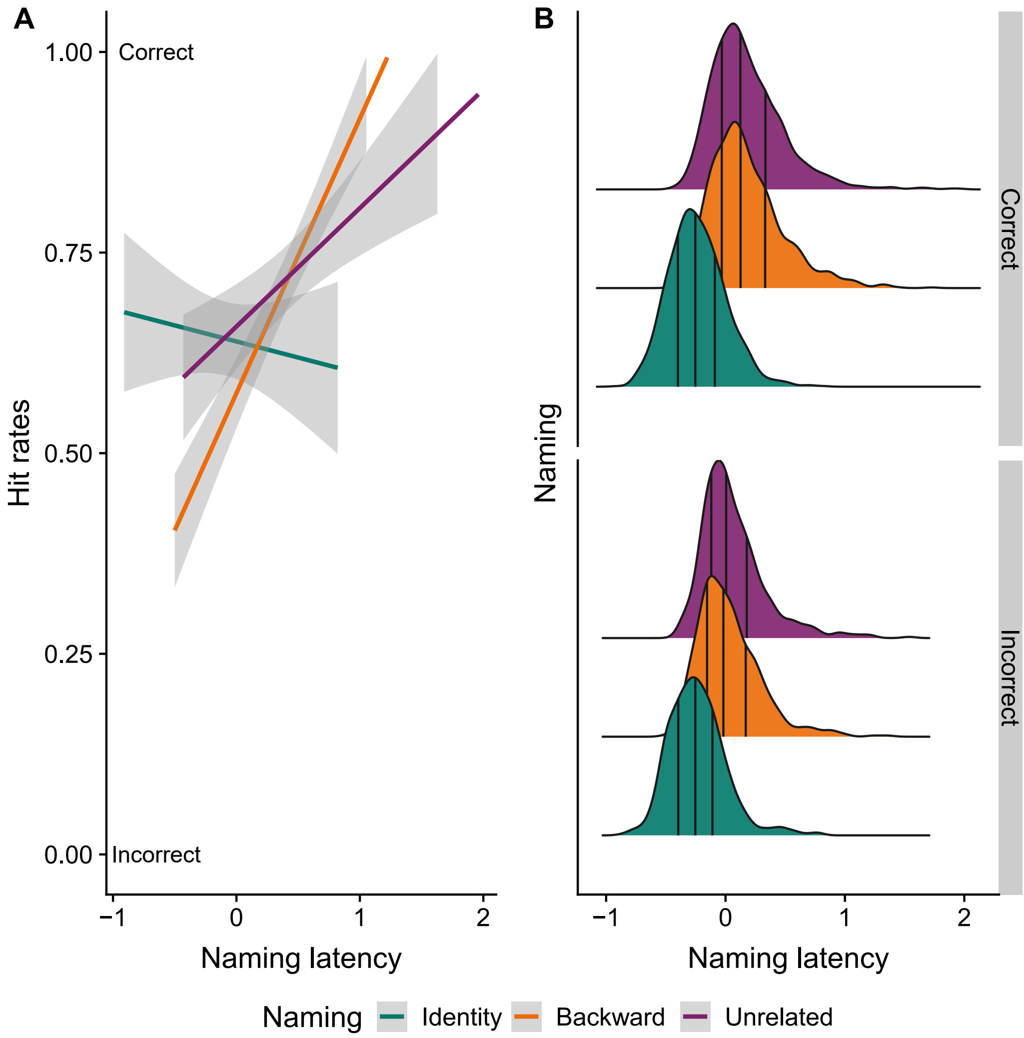
Random effects			
		Variance	SD
Participant	Intercept	1.04	1.02
	Naming latency	0.59	0.77
Item	Intercept	0.60	0.78
	Naming latency	0.70	0.84

The differences between the prime conditions led us to run two non-preregistered analyses. The first tested the combined effects of naming latency and prime condition on memory performance, with both terms entered as fixed effects. Trials on which participants hesitated or stuttered were excluded (2.7% of the data). The random effects structure included by-participant and by-item intercepts as well as by-participant and by-item slopes for naming latency.

Naming latency was a significant predictor of memory performance, with longer naming times associated with better recognition memory (see Table 3.4). There were no significant main effects of prime condition. However, both contrasts created from prime condition interacted significantly with naming latency. In the identity condition, naming latency had no effect on memory performance. As Figure 3.2 shows, these trials were remembered relatively poorly regardless of how much time was spent on naming. In contrast, in both the backward condition and the unrelated condition, longer naming latencies led to improved memory performance. A cross-over interaction between the backward condition and the unrelated condition was also observed, such that for the slowest trials, the backward condition led to better memory performance than the unrelated condition, while for the fastest trials, the unrelated condition led to better memory performance than the backward condition.

The second analysis added gaze and button-press measures to the model including naming latency and prime condition latency to examine whether

Figure 3.2: Panel A: Hit rates by prime condition and naming latency. Naming latency was binned to the second decimal point to calculate Hit rates. Panel B: stacked density plots for each prime condition of correct (top) and incorrect (bottom) responses by naming latency. The lines signify the first, second, and third quartile.



they predicted any additional variance. The gaze duration data were extracted from the time window between picture onset and the end of the trial. As seen in Table 3.5, naming latency was a better predictor of memory (β = 1.18) than either gaze duration (β = -0.03) or button-press latency (β = 0.17). Furthermore, compared to the original model, model fit did not significantly improve by adding either gaze duration (p = 0.78), or button-press latency (p = 0.36).

Table 3.5: Mixed-effects logistic regression testing the effect of naming latency, prime condition, gaze duration, and button-press latency on later memory performance (log-odds of Yes responses). All latency measures have been log-transformed and centred.

Fixed effects					
	Estimate	SE	Wald *z*	*p*	*CI*
Intercept	0.88	0.16	5.65	< 0.001	0.58, 1.20
Naming latency	1.18	0.24	4.81	< 0.001	0.70, 1.68
Gaze duration	-0.03	0.09	-0.28	0.78	-0.20, 0.15
Button-press latency	0.17	0.18	0.92	0.36	-0.19, 0.52
Identity vs. Backward & Unrelated	0.08	0.13	0.65	0.52	-0.17, 0.34
Backward vs. Unrelated	0.17	0.09	1.91	0.06	-0.01, 0.35
Naming latency:Id vs. Bw. & Un	2.29	0.41	5.65	< 0.001	1.50, 3.12
Naming latency:Bw. vs. Un	-0.68	0.33	-2.08	0.04	-1.35, -0.02

Random effects			
		Variance	SD
Participant	Intercept	1.03	1.01
	Naming latency	0.60	0.78
Item	Intercept	0.60	0.77
	Naming latency	0.71	0.84

Discussion

Our results demonstrate that naming pictures, compared to repeating their names, leads to superior recognition memory for the picture names. This is consistent with the generation effect in Zormpa, Brehm, et al. (2019). The finding that generation improves recognition memory even when picture names are used at test indicates a post-visual origin of the effect: If generation enhanced only the visual representation of the pictures, no generation effect should be found in a test using picture names. A generation effect for picture labels, as observed in this study, can arise only if the representations that generation enhances are conceptual or lexical in nature.

The study phase of the experiment was self-paced; as such, study time variations might account for the observed memory benefit for generated words. As expected, pictures were named faster when they required no generation. If longer processing time leads to more distinctive episodic representations, this would predict worse recognition for items studied for less time. Indeed, an analysis including naming latency as a predictor along with prime condition showed no main effect of prime condition. Thus, processing time accounts for substantial variance in memory performance.

However, the observed interaction between naming latency and prime condition suggests that additional time benefits memory only when it is spent preparing to name. That is, when generation was not required (identity condition), participants were poor at recognising picture names regardless of how long it had taken them to repeat them. Increasing time spent preparing to repeat a picture label does not impact memory.

In contrast, when generation was required (in the backward and unrelated conditions), longer naming time was associated with better memory. In these conditions, variations in naming time likely reflect variations in conceptual and lexical processing, showing an important link between psycholinguistic processes and memory. This claim is further supported by the exploratory analysis adding gaze duration and button-press latency to a model including naming latency: Neither explained any additional variance, indicating that it is specifically time spent preparing to name a picture that improves memory performance, not simple exposure time.

To conclude, we have replicated the generation effect in a cross-modal format, demonstrating that the generation effect observed in picture naming derives from enhanced conceptual and linguistic representations. This is important for theories of episodic memory in language, as language involves reference to objects from past contexts. Our results underscore why this might be easier for speakers than listeners—generating a name leads to a better episodic representation of the associated concept and linguistic features. This informs research on the intersection of memory and language with implications for phenomena like pronominal resolution and common ground building.

4 Answers are remembered better than the questions themselves

Abstract

A key aspect of language is how it is used in communicative contexts to identify and successfully transmit new information. New and important information can be highlighted by putting it in linguistic focus through pitch, syntactic structure, or semantic content. Previous tests of focus and memory have used structures that are rare in everyday life, overlooking the role focus plays in communication. The present study uses question-answer pairs, common in everyday conversation, to test whether questions focus their answers such that the answers are remembered better than the questions themselves. Forty-eight participants saw three pictures on the screen while listening to a recorded question-answer exchange between two people about the location of two of the displayed pictures. In a Yes/No memory test conducted online a day later, participants recognised the names of pictures that served as answers 6% more accurately than the names of pictures that appeared as questions (β = 0.27, Wald z = 4.51, 95% CI = 0.15, 0.39, p = < 0.001). Because questions and answers are, generally, uttered by two different people, the finding that answers are remembered better than questions can be used to study the memory representations developed by two interlocutors.

Introduction

We are exposed to immense amounts of language every day. For better or worse, we only retain a small amount of this verbal information. What information we do remember depends on many factors, one of which is the way that the information is presented. A robust finding from research on information structure is that when information is presented as important in a discourse, otherwise known as *focused*, it tends to be remembered better than when presented neutrally or in contrast to the focused information (Birch, Albrecht, & Myers, 2000; Birch & Garnsey, 1995; Kember, Choi, Yu, & Cutler, 2019; Mckoon, Ratcliff, Ward, & Sproat, 1993; Sturt et al., 2004; Ward & Sturt, 2007; Yang, Zhang, Wang, Chang, & Li, 2017, though c.f. Almor & Eimas, 2008). The present work investigates the effect of focus on memory via question-answer pairs, a common way to focus information in conversation.

Focus has been researched extensively in the past decades and has been associated with a number of definitions and manipulations. In terms of definition, we follow Levelt (1989) and take focus to refer to the most attended part of a discourse. Focus can be induced in many ways, including by manipulating acoustic properties (Halliday, 1967), syntax (Birch & Garnsey, 1995; Birch & Rayner, 1997), or semantic context (Cutler & Fodor, 1979).

Focused information has been associated with advantages in online processing and later memory. For instance, focused information tends to be processed for longer and in more detail (Benatar & Clifton, 2014; Birch & Garnsey, 1995; Birch & Rayner, 1997; Osaka, Nishizaki, Komori, & Osaka, 2002; Ward & Sturt, 2007) and is remembered better than information that is neutral or not focused (Birch & Garnsey, 1995; Cutler & Fodor, 1979; Fraundorf et al., 2010; Johns, Gordon, Long, & Swaab, 2014; Sturt et al., 2004, though c.f. Almor & Eimas, 2008). These phenomena have been argued to stem from a common cause: focused items are encoded more deeply than non-focused items, leading to stronger representations in the discourse model (Foraker & McElree, 2007). Indeed, Sturt et al. (2004) found that participants were more accurate at detecting changes in a text when the element that was changed had previously been focused than when it had not. Moreover, Wang, Bastiaansen, Yang, and Hagoort (2011) found larger N400 components, interpreted to reflect depth of processing, for focused items than non-focused items. Focus has also been associated with visual attention: participants spend more time reading items that are focused than items that are not focused (Benatar & Clifton, 2014; Birch & Rayner, 1997; Lowder & Gordon, 2015, though c.f. Birch & Rayner, 2010). Together, these findings suggest that focus causes people to process information more attentively and encode it more deeply, leading to stronger representations and more consistent knowledge retention.

A lot of our language experience comes from communication; as such, it is important to ensure that the effects we observe experimentally obtain when using structures that are frequent in everyday communication. However, by far the most common sentence-level manipulations of focus are clefts and pseudo-clefts (e.g. Almor & Eimas, 2008; Birch et al., 2000; Birch & Garnsey, 1995; Birch & Rayner, 1997, 2010; Foraker & McElree, 2007; Järvikivi, Pyykkönen-Klauck, Schimke, Colonna, & Hemforth, 2014; Lowder & Gordon, 2015; Morris & Folk, 1998; Sanford, Price, & Sanford, 2009), which are exceedingly rare, appearing in less than 0.1% of English sentences (Roland et al., 2007). In these structures, syntax guides attention to one element of the sentence, e.g., "It is the goat that should move next to the painting" or "What should move next to the painting is the goat".

Another way of inducing focus is the manipulation of the semantic context through questions. However, the most common manipulation in the literature tends to use indirect questions, which are still relatively infrequent in everyday communication. For example, participants might see sentences like "Everyone was wondering what would happen" or "Everyone wanted to know which item should move" and then see the answer "It turns out the goat should move next to the painting" (Benatar & Clifton, 2014; Cutler & Fodor, 1979; Sauermann, Filik, & Paterson, 2013; Sturt et al., 2004; Wang et al., 2011; Ward & Sturt, 2007; Yang et al., 2017). The answer is identical in both conditions, but the word "goat" is remembered better in the latter case.

Direct questions, like indirect ones, are very effective at eliciting focus (Chomsky, 1971). They also do so through context, such that targets (i.e., answers) are identical across conditions. They also have the benefit of being very frequent in conversation (Graesser, McMahen, & Johnson, 1994). Importantly for the present work, direct questions have been shown to have an effect on memory. Cutler and Fodor (1979) used auditory questions that put either the subject or the object of the answer in focus. In a four-alternative forced-choice (4AFC) sentence completion task, participants made fewer errors when the target response had been a focused item than a non-focused item. Yang et al. (2017) used questions that varied whether focus was placed on a word or not; at the end of a three-sentence-long narrative, that word acted as the target in a probe recognition task. Responses to those targets were faster after focusing questions than after non-focusing questions.

Research on question-answer pairs leaves unanswered whether there are memory differences between the question and the answer. Until now, researchers have always tested how a preceding question affects memory for different parts of the answer relative to each other. From a communication perspective, it is also important to know how answers are remembered relative to questions because these are uttered by different people. That is, understanding more about how questions and answers are represented gives us insight into the discourse models that different interlocutors are building. Because language production and participation in conversation can influence how items are represented, it is important to first test the effect of questions during passive comprehension. In the present study, participants hear question-answer pairs like "What should move next to the painting?" "The goat." and see pictures of a goat, a painting, and an unrelated item (a doll). We expect that the focused items (answers) will be remembered better than neutral items (questions), i.e., we expect "goat" to be remembered better than "painting".

The present study also takes a novel approach to how it measures memory: we examined the consequences of focus on memory cross-modally at a 24-hour delay. While other effects linking language and memory have been shown to last a day or even more (e.g., Gaskell, Cairney, & Rodd, 2019; Ozubko, Hourihan, & MacLeod, 2012; Zormpa, Meyer, & Brehm, 2019), the memory benefits of focus have most commonly been assessed over very short time periods. This leads to the question of whether focus truly has a lasting impact on memory and at what level the memory benefit occurs. In studies using a probe recognition task or an error detection paradigm, participants' memory is tested after two or three sentences, typically a couple of minutes after exposure (Sturt et al., 2004; Ward & Sturt, 2007; Yang et al., 2017). In Cutler and Fodor (1979), participants' memory was measured in a post-test that occurred a few minutes after exposure. One of the few studies to use a delayed study-test paradigm with a recognition memory test did find a lasting memory benefit for focused items (Fraundorf et al., 2010).

However, the comparison in that study was not between non-focused and focused items, but rather between items that received different types of focus prosodically. This could mean that the long-term memory benefit for focus does not generalise to other constructions. As such, it is not clear how large the effect of focus is on memory at a delay.

In the current study, we extend the work on the effects of focus on memory by testing whether the focus placed on answers to direct questions leads to a memory benefit. To answer this question we used a study-test recognition memory paradigm. At study, participants listened for comprehension to question-answer pairs such as "What should move next to the painting?" "The goat." while viewing three pictures (the question item: painting, the answer item: goat, and a third, unrelated item: doll). A day later, participants completed an online Yes/No memory task with the items' labels. If the type of focus elicited from direct questions leads to a memory boost for the picture names, then participants should remember the item mentioned in the answer (goat) better than the item mentioned in the question (painting). We used words in the recognition task instead of the pictures that participants saw at study to ensure that any memory benefit that is observed comes from the item labels, rather than the item pictures, localising the effect to language.

Methods

Participants

Forty-eight native Dutch speakers (38 female) aged 18-30 (M = 23) were recruited from the Max Planck Institute for Psycholinguistics participant database. They received 8 € for their participation. None disclosed any speech and language problems and all had normal or corrected-to-normal vision. We selected a sample size of 48 participants by running a power analysis in which we simulated data with effect sizes ranging from 3% to 6% memory improvement. With 128 target items (384 in total, which is as many as we could find) 42 participants would give us 80% power to detect condition-level differences of 4% or greater. We tested 48 participants to have a balanced number of participants in each list. Ethical approval to conduct this study was given by the Ethics Board of the Social Sciences Faculty of the Radboud University.

Materials

Pictures.

In the first phase of the experiment, 384 colour photographs were used as stimuli. Most (N = 322) were sourced from the BOSS picture database (Brodeur et al., 2010, 2014), but 62 came from other stimulus sets (Brady, Konkle, Alvarez, & Oliva, 2008, 2013; Moreno-Martínez & Montoro, 2012), or Wikimedia Commons. A full list of the stimuli and their sources can be found in the Appendix.

All pictures were normed for name agreement, familiarity, visual complexity (measured in JPEG size, see Machado et al., 2015), $\log_{10}$ frequency, and length (measured in letters). This was done in stages. First, 387 pictures were normed for name agreement by 15 participants recruited from the Max Planck Institute for Psycholinguistics participant database in an online study. Pictures with less than 83% name agreement were replaced with pictures from the BS database that were previously normed in Dutch by Decuyper et al. (in prep). Familiarity norms for all pictures from the BOSS set were drawn from Brodeur et al. (2010, 2014), and the remaining pictures were normed for familiarity by 8 native Dutch speakers employed at the

Max Planck Institute for Psycholinguistics. Estimates of $\log_{10}$ frequency for all items were taken from the SUBTLEX-NL corpus (Keuleers et al., 2010). These measures were used to split pictures in two balanced sets A and B using Match (van Casteren & Davis, 2007): name agreement (M_A = 0.93, M_B = 0.93), familiarity (M_A = 4.27, M_B = 4.26), visual complexity (M_A = 48660, M_B = 48310), $\log_{10}$ word frequency (M_A = 2.21, M_B = 2.26) and length (M_A = 6.72, M_B = 6.83). These sets were counterbalanced across four lists such that in two lists, set A was used as targets and set B as foils, and in two lists, set A was used as foils and set B as targets.

Within sets A and B, three further subgroups were created using Match resulting in subsets A1, A2, A3, and B1, B2, and B3. These were used to assign pictures to the question, answer, and unmentioned conditions used in the study phase of the experiment. Subsets A3 and B3 were always used as unmentioned items and the question and answer conditions were assigned to subsets A1 and A2 or B1 and B2 across four counterbalanced lists. The three subsets were combined into trials pseudo-randomly with one item from each such that none of the three pictures were semantically related or started with the same phoneme.

Pictures were presented in 300x300 pixel resolution.

Study phase.

In each trial in the study phase, participants saw three pictures and heard a conversation snippet between two native Dutch speakers (one female and one male). Speakers were recorded using Shure SM10A microphones while participating in a version of the experiment in which two participants asked and answered questions about the position of objects on the screen. This was done to preserve as natural an intonation as possible. The recordings were then edited to remove static and normalised in volume using Audacity (version 2.0.6; Audacity Team, 2014). Silences at the end of the recordings were removed, and silences at the beginning of recordings that were relatively long or short compared to the others were shortened or lengthened accordingly such that trials began with 1132 ms silence on average (*SD* 255).

Test phase.

In the test phase, participants saw the most common Dutch name for each of the 384 pictures. These were presented one at a time centrally on a white background (RGB: 255,255,255) in Calibri font, size 45.

Procedure

In the study phase, participants were tested one or two at a time in an experiment room with two soundproof booths in a session that lasted approximately 25 minutes. The experiment was controlled using Presentation (version 18.3; Neurobehavioral Systems, Berkeley, CA, USA) and was displayed on a 24" monitor (1920x1080 pixel resolution). First, participants completed four practice trials for which they received feedback and were allowed to ask questions. They then completed 72 experimental trials for which no feedback was provided. All trials started with a fixation cross displayed in the middle of the screen for 500 ms, followed by a blank screen that appeared for another 500 ms. Then, participants saw the three pictures, each occupying one of the four corners of the screen. The position of the each item role (question, answer, unmentioned, empty) was counterbalanced within lists and all combinations of role and location were used 4 times per list. The trials began with a silent period (duration 565-2448 ms, M = 1132 ms), followed by the conversation snippet (duration 2460-4052 ms, M = 2981 ms). Participants then pressed on the space bar to move to the next trial.

Of the 72 experimental trials, eight were catch trials. On these catch trials, participants were given a comprehension question after pressing the space bar to end the trial. The question queried the location of one of the item roles (question, answer, unmentioned, empty). Participants had to answer the question by selecting one of the four corners using the keyboard. No feedback was provided and each role and each location was queried twice. Two out of the four practice trials were catch trials in which participants received feedback on whether they answered the question correctly.

The second phase of the experiment was conducted online the following day. Participants were sent a link to an online Yes/No recognition memory test and were given 8 hours to complete it. The names of all pictures that were shown the day before were presented, mixed with an equal number of foils. Participants were instructed to respond with "Yes" to the names of all pictures seen the previous day, including those that were unmentioned. There was no time limit for the second session, but it usually lasted 10-15 minutes.

Analysis

Preregistered exclusion criteria included failure to complete the second phase of the experiment or below chance performance in either phase of the experiment (under 25% of the catch trials in the first phase or under 50% of Hits in the memory task for the items in the questions and answers). No participants were excluded by these criteria.

Analyses were run using the lme4 package (version 1.1-21; Bates et al., 2015) in R (version 3.5.3; R Core Team, 2019) with the optimiser BOBYQA (Powell, 2009). The dependent variable were the log-odds of Yes responses in the memory task. The random effects structure included random intercepts for items and participants, as well as any random slopes licensed for each random intercept. The exact structure was determined in a data-driven way starting from the maximal model and eliminating the slopes that explained the least variance if the model did not converge, or slopes that were correlated at a level of 0.95 or above with the random intercept. We report the final random effects structure in each model table.

Two preregistered analyses were run. The first aimed to ensure that participants were able to discriminate old from new items and used the target vs. foil contrast as the only fixed effect. This was sum-to-zero contrast coded with targets coded as 0.5 and foils as -0.5. Next, the primary analysis tested the hypothesis that answers to questions are remembered better than questions. In this analysis, the answer vs. question fixed effect was also sum-to-zero contrast coded with answers coded as 0.5 and questions as -0.5.The preregistration can be found on the OSF project for this study (https://osf.io/x45ad/registrations).

Results

Accuracy in the catch trials at study was high (*M* = 91%, *SD* = 14%), meaning that participants were paying attention during the comprehension task. However, accuracy in the memory task was lower (*M* = 58%, *SD* = 49%). Performance for each item role in the memory task can be seen in Figure 4.1. Participants were generally conservative: They were very good at (correctly) rejecting new items (19% false alarm rate), but they also (falsely) rejected many of the old items. Despite the relatively low accuracy, the pattern of responses to each condition follows the predicted pattern: Recognition of old items was best for the answers and worst for the unmentioned items.

Table 4.1: Mixed-effects logistic regression testing the effect of Probe, i.e., targets vs. foils.

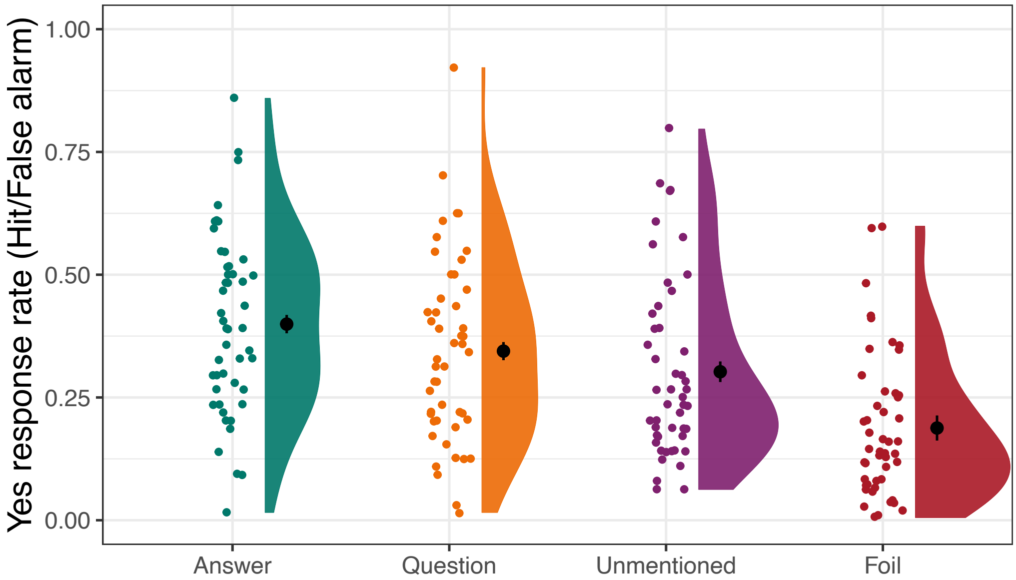
Fixed effects					
	Estimate	SE	Wald z	p	*CI*
Intercept	-1.31	0.14	-9.12	< 0.001	-1.59, -1.02
Targets vs. Foils	1.13	0.10	11.35	< 0.001	0.94, 1.34

Random effects			
		Variance	SD
Participant	Intercept	0.94	0.97
	Targets vs. Foils	0.33	0.58
Item	Intercept	0.12	0.35
	Targets vs. Foils	0.37	0.61

The first analysis examined the effect of probe, i.e., whether participants could successfully distinguish old items (targets) from new ones (foils). The full logistic regression model for the analysis comparing targets and foils can be seen in Table 4.1. The random structure included by-participant and by-item intercepts and slopes for the target vs. foil contrast (maximal model). The negative intercept shows that participants had a large "No" bias in this experiment. The positive estimate for the target vs. foil comparison shows that participants responded positively to targets more often than to foils, i.e. they were able to reliably distinguish old and new items at test.

The second analysis tested the effect of focus to examine if answers were indeed remembered better than questions. This analysis was run on the memory trials that included items that were verbally mentioned at study, i.e. the items that were used in the questions and the answers. The full logistic regression model for this analysis is displayed in Table 4.2. The random structure included by-participant and by-item intercepts and by-item random slopes for the answer vs. question contrast. Again, the negative intercept shows participants had a bias towards responding "No" overall. The positive estimate for the answer vs. question contrast shows that par-

Figure 4.1: "Yes" responses to each item role in the memory task.



ticipants responded with "Yes" to answers more often than to questions. In other words, participants were more accurate when recognising items that appeared as answers than items that appeared as questions.

Table 4.2: Mixed-effects logistic regression testing the effects of Focus, i.e., answers vs. questions.

Fixed effects					
	Estimate	SE	Wald z	p	*CI*
Intercept	−0.63	0.14	−4.53	< 0.001	−0.91, −0.39
Answers vs. Questions	0.27	0.06	4.51	< 0.001	0.15, 0.39

Random effects			
		Variance	SD
Participant	Intercept	0.87	0.93
Item	Intercept	0.89	0.30
	Answers vs. Questions	0.07	0.27

Discussion

In this experiment we were interested in whether the answers to questions are remembered better than the questions themselves. We predicted that memory for answers would be stronger than for questions, because answers to questions are, by definition, focused. Our results confirmed this prediction: Accuracy in the memory task increased by 6% when a word had been used in an answer compared to when it had appeared in a question. This increase was observed despite the simplicity of the dialogues, which involved no protagonists or narrative, and despite the change of modality of the items at study (presented as auditory words and pictures) versus at test (presented as written words). This means that the effect was driven by focus at the level of mental representations of words. The clear advantage for answers is important for the study of memory representations that are developed during communication, as it demonstrates that people are effective in drawing their interlocutors' attention (or in this case a passive listener's attention) to the items they consider important. These results are consistent with earlier work showing answers have a benefit over other items in a sentence, but extends it by showing that answers also have a benefit over the questions.

An unexpected finding of this study was that memory for all items was relatively poor. There are many potential reasons for this, including the relatively artificial nature of the dialogues, the fact that participants merely observed a conversation without participating in it, the inclusion of unmentioned items in the memory test, the modality shift between study and test, and the long lag between study and test. While we cannot assess the impact of each of these variables, we believe the most likely cause for low performance was the inclusion of unmentioned items, which were probably not well-encoded, in the memory test. Unmentioned items were largely rejected in this study and, in addition to the foil items which require a "No" answer, they may have skewed our participants' criterion to elicit a "No" bias in responding. The modality shift and long interval (20 to 28 hours) between study and test may have exacerbated this problem, though we note that the same procedure has been successfully used to test memory in language production studies (Zormpa, Meyer, & Brehm, 2019), so we consider this less likely.

The observed low performance on unmentioned items is also consistent with the role of attention in memory. The memory advantage for answers over questions, along with the other processing benefits ascribed to focus, has been theorised to stem from the attention afforded to focused items. When items are attended to more, they are encoded more deeply and with more detail, which leads to stronger memory representations. Research using eye-tracking has indeed found some evidence for increased visual attention to focused items (Benatar & Clifton, 2014; Birch & Rayner, 1997; Lowder & Gordon, 2015, though c.f. Birch & Rayner, 2010). From a communicative perspective, when someone asks a question they presumably do so because they are interested in the answer. As such, the memory benefit for answers is fully consistent with the role of attention for focused items. It is reasonable to assume that a person attends more to the answer than to the question; future work might assess how much attention is allocated to each using more sensitive online methodologies.

Studying focus in the context of communication is important, as a significant proportion of our linguistic experience comes from communicating with others. Question-answer pairs are useful for researching focus in conversations for two reasons. Firstly, question-answer pairs are quite frequent in conversation. In a corpus of spontaneous conversations between native speakers of German (GECO; Schweitzer & Lewandowski, 2013, 2014), we see that 845 of 2689 conversational turns (31%) are direct questions. In contrast, only 5 of 2689 turns were clefts (about 0.002%) and many of these clefts contained grammatical errors or were otherwise syntactically ill-formed. The low frequency of clefts and pseudoclefts is along the same order of magnitude as estimates from English in corpora of written and spoken language (Roland et al., 2007). This frequency difference is quite remarkable and matches our own frequent observations of questions and infrequent observations of clefts and pseudo-clefts in everyday conversation. This makes question-answer pairs more representative of how focus is used in natural communication.

In addition, because question-answer pairs are more common in conversation, this opens the door to further examinations of how focus affects information transfer between people in simple scenarios that are more reflective of spoken language use. By using questions and answers, researchers can easily extend this research to examine the effect of focus in conversation in general, such as how it differs across different interlocutor roles or with conversational engagement. We highlight the utility of simple manipulations in studies of language and memory that better reflect spoken language, and suggest that future experimental paradigms might be considered in light of the types of structures frequently used in spoken as opposed to written language.

To conclude, this work has shown that questions effectively focus their answers and, as a result, the answers to the questions are remembered better than the questions themselves. This finding is important for the study of memory representations in conversation because it allows for the exploration of the effect of focus in guiding attention in a conversation. The discourse function of answers likely allows us to preferentially remember new information, which might be a strategic choice given the amount of language we are exposed to each day.

Appendix

Below are the stimuli sets in the present and the following chapter. Agr. refers to naming agreement rate, L refers to length (measured) in letters, Fam refers to familiarity, VC refers to visual complexity (measured in file size), WF refers to word frequency ($\log_{10}$), and Db refers to Database or the source of the images. Most of the images come from the BOSS database (Brodeur et al., 2010, 2014) (BS; 322). Additional pictures come from Moreno-Martínez and Montoro (2012) (MM; 26) and from Wikimedia Commons (WC; 26). The rest come from the UCSD Vision and Memory lab stimuli resources: from Brady et al. (2008) (B08; 4) and from the exemplar section of Brady et al. (2013) (B13; 2), from the Pixabay website (PB;3) and from the Maxpixel website (MP;1).

ID	Filename	Dutch	Agr	L	Fam	VC	WF	Db
01	accordion01.jpg	accordeon	0.73	9	4.26	70908	1.72	BS
02	acousticguitar02.jpg	gitaar	1.00	6	4.52	37099	2.70	BS
03	africanelephant.jpg	olifant	1.00	7	4.48	54268	2.72	BS
04	almond.jpg	amandelen	0.87	9	4.30	64284	1.86	BS
05	aluminiumfoil.jpg	aluminiumfolie	0.80	14	4.79	37394	1.42	BS
06	anchor.jpg	anker	1.00	5	4.20	42072	2.34	BS
07	ant.jpg	mier	0.93	4	4.62	43698	2.05	BS
08	antlers.jpg	gewei	0.94	5	4.02	43056	1.72	BS
10	apple07.jpg	appel	1.00	5	4.60	54331	2.65	BS
12	apron.jpg	schort	1.00	6	4.57	60088	2.09	BS
13	aquarium.jpg	aquarium	1.00	8	4.29	60477	2.10	BS
14	arm.jpg	arm	0.93	3	4.93	28827	3.54	BS
15	armchair02.jpg	stoel	0.87	5	4.62	63967	3.35	BS
16	arrow02.jpg	pijl	0.73	4	4.24	35143	2.49	BS

18	asparagus.jpg	asperges	0.73	8	4.00	43074	1.59	BS
19	atm.jpg	pinautomaat	0.87	11	4.71	50180	1.60	BS
20	avocado01.jpg	avocado	1.00	7	4.30	60193	1.38	BS
21	axe01.jpg	bijl	0.93	4	3.70	34824	2.61	BS
23	balloon01b.jpg	ballon	0.93	6	4.40	50674	2.37	BS
24	banana01.jpg	banaan	1.00	6	4.70	32920	2.37	BS
25	bandaid01.jpg	pleister	0.93	8	3.90	45594	2.03	BS
26	barbedwire01.jpg	prikkeldraad	0.95	12	4.05	79475	2.11	BS
28	barrel01.jpg	ton	0.93	3	4.14	56318	2.78	BS
29	basketball01.jpg	basketbal	0.87	9	4.79	85071	2.59	BS
32	battery02b.jpg	batterij	1.00	8	4.60	49647	2.47	BS
34	bed.jpg	bed	1.00	3	4.80	45879	4.02	BS
35	belt.jpg	riem	1.00	4	4.40	35757	2.79	BS
36	bib.jpg	slabbetje	0.93	9	3.80	83348	1.23	BS
37	frisbee.jpg	frisbee	0.95	7	4.40	51568	1.91	BS
38	bikepump01.jpg	fietspomp	0.93	9	4.52	33233	0.70	BS
39	binoculars01b.jpg	verrekijker	1.00	11	4.00	53252	2.09	BS
40	bleachers.jpg	tribune	0.92	7	4.43	58330	2.08	BS
42	book01b.jpg	boek	0.93	4	4.50	35040	3.82	BS
43	boot02b.jpg	laars	1.00	5	4.40	49740	2.26	BS
44	bow.jpg	boog	1.00	4	4.20	43029	2.57	BS
45	bowlingball.jpg	bowlingbal	1.00	10	4.43	64764	1.34	BS
46	bowlingpin.jpg	kegel	0.80	5	4.46	30315	1.30	BS
47	bowrake.jpg	hark	0.97	4	4.48	34167	1.83	BS
48	bowtie.jpg	strik	0.87	5	4.55	46424	1.95	BS
49	bracelet01.jpg	armband	0.93	7	3.40	47155	2.44	BS
52	bridge.jpg	brug	0.93	4	4.52	36709	3.29	BS
53	broccoli01a.jpg	broccoli	1.00	8	4.70	51248	2.07	BS

54	broom01.jpg	bezem	1.00	5	4.30	37083	2.22	BS
55	bucket01a.jpg	emmer	0.93	5	3.90	52470	2.47	BS
56	bull.jpg	stier	0.92	5	4.27	49807	2.62	BS
57	bullet.jpg	kogel	0.93	5	4.07	42743	3.28	BS
59	bus.jpg	bus	0.80	3	4.69	62965	3.45	BS
60	button01.jpg	knoop	1.00	5	4.40	45276	2.62	BS
61	cactus.jpg	cactus	1.00	6	4.14	55452	1.90	BS
62	calculator01.jpg	rekenmachine	0.93	12	4.30	48238	1.42	BS
63	calendar.jpg	kalender	1.00	8	4.74	49586	2.17	BS
64	candle08b.jpg	kaars	1.00	5	4.30	38590	2.41	BS
65	cane.jpg	wandelstok	0.67	10	3.50	33439	1.54	BS
66	cannon.jpg	kanon	1.00	5	4.26	47925	2.43	BS
68	car.jpg	auto	1.00	4	4.57	59500	4.30	BS
69	carrot01.jpg	wortel	0.87	6	4.40	42292	2.43	BS
70	cashregister01.jpg	kassa	1.00	5	4.48	52974	2.52	BS
71	cat.jpg	kat	0.92	3	4.48	67584	3.36	BS
74	cherry01.jpg	kersen	0.93	6	4.67	49217	2.02	BS
76	chimney.jpg	schoorsteen	1.00	11	4.38	66291	2.28	BS
78	cigar.jpg	sigaar	0.93	6	4.12	42249	2.63	BS
80	clothespin03b.jpg	wasknijper	0.93	10	4.30	37415	0.30	BS
81	cloud.jpg	wolk	1.00	4	4.74	36117	2.38	BS
83	coatrack.jpg	kapstok	1.00	7	4.55	30473	1.51	BS
84	cobra.jpg	slang	1.00	5	4.05	45781	2.98	BS
85	coconut.jpg	kokosnoot	0.93	9	3.90	93013	1.88	BS
86	coffeebean.jpg	koffiebonen	0.93	11	4.30	61186	1.20	BS
87	comb02a.jpg	kam	0.93	3	4.40	50754	2.37	BS
88	computerkeyboard02.jpg	toetsenbord	0.87	11	4.70	51788	1.69	BS
89	computermouse06.jpg	muis	0.73	4	4.80	44456	2.69	BS

90	cookie01.jpg	koekje	1.00	6	4.40	58403	2.55	BS
91	cork02.jpg	kurk	0.93	4	4.33	64307	1.86	BS
92	corkboard.jpg	prikbord	1.00	8	4.52	75407	1.45	BS
93	cow.jpg	koe	1.00	3	4.71	45178	2.91	BS
94	crab01.jpg	krab	0.87	4	4.12	47554	2.23	BS
95	crocodile.jpg	krokodil	0.87	8	4.07	33126	2.33	BS
96	croissant01.jpg	croissant	1.00	9	4.50	64233	1.42	BS
97	cross01.jpg	kruis	0.92	5	4.55	35582	2.96	BS
98	crown.jpg	kroon	1.00	5	4.43	81444	2.80	BS
99	cd.jpg	cd	0.83	2	4.70	61627	2.58	BS
101	curtain.jpg	gordijn	1.00	7	4.76	46504	2.29	BS
102	daddylonglegs.jpg	spin	1.00	4	4.33	37865	2.53	BS
103	dartboard.jpg	dartbord	0.97	8	4.46	112243	0.70	BS
104	dice05a.jpg	dobbelsteen	1.00	11	4.50	46212	1.51	BS
105	discoball.jpg	discobal	0.86	8	4.55	99149	0.70	BS
106	dishsoap.jpg	afwasmiddel	1.00	11	4.40	36080	1.04	BS
108	doghouse.jpg	hondenhok	0.93	9	4.50	52350	1.54	BS
110	dolphin01.jpg	dolfijn	1.00	7	4.48	42771	1.92	BS
111	donut.jpg	donut	1.00	5	4.74	46832	2.29	BS
113	doorhandle.jpg	deurklink	1.00	9	4.78	51170	0.95	BS
114	doorlock.jpg	slot	0.89	4	3.90	65230	3.36	BS
115	dreamcatcher.jpg	dromenvanger	0.92	12	4.05	49194	0.85	BS
117	drumset.jpg	drumstel	1.00	8	4.71	64091	1.83	BS
118	duck.jpg	eend	1.00	4	4.50	44603	2.60	BS
119	ear.jpg	oor	0.93	3	4.95	38691	3.04	BS
120	eggplant.jpg	aubergine	1.00	9	4.00	34263	1.49	BS
121	elbow.jpg	elleboog	0.93	8	4.90	36146	2.16	BS
122	endive.jpg	witlof	0.87	6	3.70	41816	0.78	BS

123	englishcucumber.jpg	komkommer	1.00	9	4.74	32059	1.77	BS
124	envelope03a.jpg	envelop	0.93	7	4.60	32026	2.44	BS
125	eraser.jpg	gum	1.00	3	4.30	43857	1.34	BS
126	escalator.jpg	roltrap	0.93	7	4.81	51583	1.61	BS
127	eye.jpg	oog	1.00	3	4.90	49189	3.48	BS
128	fan.jpg	ventilator	0.80	10	4.20	53193	2.05	BS
129	faucet.jpg	kraan	0.95	5	4.71	53354	2.45	BS
130	feather03a.jpg	veer	1.00	4	3.90	31960	2.18	BS
131	fence02.jpg	hek	0.93	3	4.38	52813	3.00	BS
132	fingerprint.jpg	vingerafdruk	0.93	12	4.69	70570	2.28	BS
133	flag.jpg	vlag	0.93	4	4.21	35789	2.89	BS
134	flamingo.jpg	flamingo	1.00	8	4.43	38792	1.83	BS
135	flashlight02b.jpg	zaklamp	0.93	7	4.30	34581	2.35	BS
136	foot.jpg	voet	0.73	4	4.81	33783	3.35	BS
138	fork03c.jpg	vork	1.00	4	4.60	36600	2.36	BS
139	frenchfries.jpg	friet	0.93	5	4.93	49680	2.17	BS
140	fridge.jpg	koelkast	0.93	8	4.88	37290	2.81	BS
141	funnel.jpg	trechter	1.00	8	4.63	33083	1.38	BS
142	garbagecan02.jpg	prullenbak	0.73	10	4.57	44869	1.79	BS
143	garlic01a.jpg	knoflook	1.00	8	4.60	52212	2.29	BS
144	gift01.jpg	cadeau	1.00	6	4.71	64931	3.11	BS
146	giraffe.jpg	giraffe	1.00	7	4.43	47177	1.75	BS
147	glass02a.jpg	beker	0.87	5	4.60	35941	2.59	BS
148	glasses01a.jpg	bril	1.00	4	4.30	46926	3.03	BS
151	granolabar01.jpg	mueslireep	0.92	10	4.74	70862	0.48	BS
153	grater01a.jpg	rasp	0.80	4	4.30	55773	0.85	BS
154	greatwhiteshark.jpg	haai	0.93	4	4.33	36493	2.62	BS
155	greywolf.jpg	wolf	0.95	4	4.24	59184	2.95	BS

156	grizzly.jpg	beer	0.89	4	4.40	88737	3.05	BS
158	hairdryer02a.jpg	föhn	0.87	4	4.20	39008	1.38	BS
160	hammer01.jpg	hamer	1.00	5	4.20	33757	2.57	BS
161	hand01b.jpg	hand	1.00	4	4.93	42833	3.94	BS
162	handcuffs.jpg	handboeien	0.93	10	4.48	60747	2.66	BS
163	handfan01b.jpg	waaier	0.86	6	3.70	69687	1.66	BS
164	hanger02a.jpg	kleerhanger	0.73	11	4.50	44033	1.30	BS
166	headphones02b.jpg	koptelefoon	0.93	11	4.40	51816	1.96	BS
167	helicopter.jpg	helikopter	1.00	10	4.24	40784	2.98	BS
168	helmet.jpg	helm	0.80	4	3.57	46896	2.69	BS
169	hen.jpg	kip	1.00	3	4.43	55201	3.22	BS
170	hinge.jpg	scharnier	0.93	9	3.40	55660	1.11	BS
171	hippopotamus.jpg	nijlpaard	0.92	9	4.24	53136	1.87	BS
172	horse.jpg	paard	1.00	5	4.45	46055	3.56	BS
173	horseshoe.jpg	hoefijzer	0.92	9	3.98	59722	1.51	BS
174	hourglass.jpg	zandloper	1.00	9	3.80	35630	1.72	BS
175	humanskeleton.jpg	skelet	1.00	6	4.71	46632	2.10	BS
176	humanskull.jpg	schedel	0.80	7	4.43	57294	2.80	BS
177	icecreamcone01a.jpg	ijshoorntje	0.67	11	4.20	40215	0.48	BS
178	iceskate.jpg	schaats	0.73	7	4.00	51037	1.46	BS
179	iron01b.jpg	strijkijzer	0.73	11	4.40	43835	1.51	BS
180	ironingboard01.jpg	strijkplank	0.80	11	4.52	38894	0.95	BS
181	jackrabbit.jpg	konijn	0.73	6	4.43	63998	2.92	BS
182	jar03.jpg	pot	0.80	3	3.90	40081	3.13	BS
183	jellyfish.jpg	kwal	0.97	4	3.71	42222	1.90	BS
184	kangaroo.jpg	kangoeroe	1.00	9	4.21	41275	1.84	BS
185	key01.jpg	sleutel	1.00	7	4.60	45661	3.55	BS
186	kite.jpg	vlieger	1.00	7	4.43	40450	2.22	BS

187	kiwi03.jpg	kiwi	1.00	4	4.71	50947	1.46	BS
188	knee.jpg	knie	1.00	4	4.93	37714	2.65	BS
189	knife03.jpg	mes	0.93	3	4.81	27852	3.31	BS
190	ladder.jpg	ladder	1.00	6	4.60	35927	2.63	BS
191	ladybug03.jpg	lieveheersbeestje	0.93	17	4.69	45823	1.28	BS
193	laptop01a.jpg	laptop	0.93	6	4.60	44118	2.41	BS
194	laundrybasket01a.jpg	wasmand	1.00	7	4.62	52681	1.60	BS
195	lawnmower.jpg	grasmaaier	0.92	10	4.55	52207	1.91	BS
196	leaf02a.jpg	blad	0.93	4	4.40	33676	2.69	BS
197	leek.jpg	prei	1.00	4	3.80	38202	1.11	BS
198	lemon02.jpg	citroen	1.00	7	4.60	38972	2.36	BS
199	lettuce.jpg	sla	0.87	3	4.60	62503	3.56	BS
200	lifejacket.jpg	reddingsvest	0.73	12	4.00	65499	1.69	BS
201	lighter01.jpg	aansteker	1.00	9	4.20	39461	2.40	BS
202	lighthouse.jpg	vuurtoren	1.00	9	4.10	39450	2.17	BS
203	lion.jpg	leeuw	1.00	5	4.40	47623	2.81	BS
204	lipstick02a.jpg	lippenstift	0.73	11	4.10	29513	2.45	BS
205	lollipop01.jpg	lolly	1.00	5	4.76	34652	1.86	BS
207	magneticcompass.jpg	kompas	0.95	6	4.31	68714	2.30	BS
208	magnifyingglass01b.jpg	vergrootglas	0.80	12	3.83	34601	1.53	BS
209	mailbox02.jpg	brievenbus	0.93	10	4.69	36569	2.27	BS
210	marble.jpg	knikker	0.94	7	4.32	52109	1.79	BS
211	mascarabrush.jpg	mascara	0.87	7	4.20	28738	1.76	BS
212	masquerademask01.jpg	masker	0.89	6	4.17	53166	2.93	BS
213	match.jpg	lucifer	1.00	7	4.40	27371	2.45	BS
214	mattress.jpg	matras	1.00	6	4.86	42676	2.35	BS
215	medal02b.jpg	medaille	1.00	8	3.80	60975	2.65	BS
216	microphone01.jpg	microfoon	1.00	9	4.60	40963	2.66	BS

217	microscope.jpg	microscoop	0.87	10	3.80	40648	1.93	BS
218	microwave.jpg	magnetron	1.00	9	4.69	43203	2.18	BS
219	mirror02.jpg	spiegel	1.00	7	4.52	39928	3.08	BS
221	monarchbutterfly.jpg	vlinder	1.00	7	4.62	82108	2.43	BS
222	moon.jpg	maan	0.87	4	4.37	48119	3.27	BS
224	mousetrap.jpg	muizenval	0.97	9	3.30	51992	1.52	BS
225	mug05.jpg	mok	0.93	3	4.88	42521	1.70	BS
226	mushroom01.jpg	champignon	1.00	10	4.60	45081	1.15	BS
227	nailclipper03b.jpg	nagelknipper	0.87	12	4.40	48971	0.85	BS
228	nailpolish03b.jpg	nagellak	1.00	8	4.10	42763	1.87	BS
229	necklace.jpg	ketting	1.00	7	4.20	35994	2.92	BS
233	onion.jpg	ui	1.00	2	4.60	48034	2.01	BS
234	orange.jpg	sinaasappel	0.86	11	4.70	73524	1.89	BS
235	ostrich.jpg	struisvogel	1.00	11	3.98	41991	1.23	BS
236	pacifier02d.jpg	speen	0.93	5	3.80	49650	1.23	BS
237	paintbrush01.jpg	kwast	0.80	5	4.10	33119	1.88	BS
239	panda.jpg	panda	0.73	5	4.55	45433	1.63	BS
240	paperclip03.jpg	paperclip	0.87	9	4.50	48584	1.46	BS
241	diaper01c.jpg	luier	0.92	5	3.56	55791	2.21	BS
242	parkfountain.jpg	fontein	0.97	7	4.43	67290	2.24	BS
243	parrot01.jpg	papegaai	0.93	8	4.10	49518	2.16	BS
244	peacock.jpg	pauw	1.00	4	4.19	63447	1.63	BS
245	peanut01.jpg	pinda	1.00	5	4.20	44515	1.97	BS
246	pear01.jpg	peer	1.00	4	4.50	40784	1.94	BS
247	pen04b.jpg	pen	0.93	3	4.80	32045	2.98	BS
248	pencil01.jpg	potlood	1.00	7	4.70	31695	2.38	BS
249	pencilsharpener02a.jpg	puntenslijper	0.87	13	4.30	55869	0.30	BS
251	pepper04a.jpg	paprika	1.00	7	4.60	47323	1.77	BS

252	perfume01a.jpg	parfum	0.87	6	4.20	50340	2.68	BS
253	photocopier.jpg	printer	0.67	7	4.52	48296	1.60	BS
254	pickle01a.jpg	augurk	0.94	6	4.26	58124	1.78	BS
255	pig.jpg	varken	1.00	6	4.36	46207	3.03	BS
256	pigeon.jpg	duif	1.00	4	4.50	44327	2.37	BS
257	pill.jpg	pil	0.80	3	4.00	42008	2.61	BS
258	pillow01a.jpg	kussen	1.00	6	4.40	41659	3.30	BS
259	pineapple01a.jpg	ananas	1.00	6	4.50	65635	2.05	BS
260	pizza.jpg	pizza	1.00	5	4.40	79220	3.03	BS
262	plate01b.jpg	bord	0.93	4	4.60	27927	3.08	BS
264	potato02b.jpg	aardappel	1.00	9	4.50	54525	2.17	BS
265	pumpkin.jpg	pompoen	1.00	7	4.71	58482	2.04	BS
266	puzzlepiece.jpg	puzzelstuk	1.00	10	4.30	51951	0.30	BS
269	radiator.jpg	verwarming	0.67	10	4.38	72251	2.34	BS
270	raspberry01.jpg	framboos	0.87	8	4.70	50047	1.11	BS
271	razor01.jpg	scheermes	1.00	9	4.30	37721	2.14	BS
272	redfox.jpg	vos	0.97	3	4.24	68340	2.52	BS
273	remotecontrol04.jpg	afstandsbediening	0.87	17	4.40	39627	2.42	BS
274	rhinoceros02.jpg	neushoorn	1.00	9	4.29	59170	2.04	BS
275	rice.jpg	rijst	1.00	5	4.50	41025	2.60	BS
276	ring01.jpg	ring	1.00	4	4.30	46958	3.36	BS
277	road02.jpg	weg	0.87	3	4.79	38963	4.81	BS
278	rock01a.jpg	steen	0.93	5	4.10	60084	3.19	BS
279	rollingpin01a.jpg	deegroller	0.93	10	3.80	30742	0.70	BS
280	rope03.jpg	touw	1.00	4	3.90	46827	3.06	BS
281	rose.jpg	roos	1.00	4	4.71	43371	2.71	BS
282	ruins.jpg	ruïnes	0.73	6	3.49	54599	2.01	BS
283	ruler04.jpg	liniaal	0.93	7	4.40	28977	1.40	BS

284	safe.jpg	kluis	1.00	5	4.26	37083	3.19	BS
285	safetypin.jpg	veiligheidsspeld	0.67	16	4.20	33952	1.04	BS
286	saltshaker03a.jpg	zout	0.73	4	4.76	37247	2.83	BS
287	sandal.jpg	sandaal	1.00	7	4.30	61789	1.04	BS
288	sandcastle.jpg	zandkasteel	1.00	11	4.36	58132	1.11	BS
289	sausage.jpg	worst	1.00	5	4.24	38992	2.59	BS
290	saw02b.jpg	zaag	1.00	4	3.70	33996	2.19	BS
291	saxophone.jpg	saxofoon	0.93	8	4.19	52052	1.64	BS
292	scale01a.jpg	weegschaal	1.00	10	4.10	44620	1.91	BS
293	scarf.jpg	sjaal	1.00	5	4.50	46601	2.37	BS
294	scissors01.jpg	schaar	1.00	6	4.50	31110	2.45	BS
295	scooter.jpg	step	1.00	4	4.43	39405	1.38	BS
296	scorpion.jpg	schorpioen	0.87	10	3.93	42007	2.13	BS
297	screwdriver04b.jpg	schroevendraaier	1.00	16	4.10	33785	1.99	BS
298	seal.jpg	zeehond	1.00	7	4.36	35679	1.42	BS
299	seashell01.jpg	schelp	0.86	6	4.00	57100	1.78	BS
300	sewingmachine01a.jpg	naaimachine	1.00	11	3.90	46416	1.26	BS
301	sheep.jpg	schaap	1.00	6	4.43	69426	2.46	BS
302	shoelace.jpg	veter	0.87	5	4.30	37998	1.77	BS
303	shoppingcart.jpg	winkelwagen	0.87	11	4.71	100548	1.20	BS
304	shoulder.jpg	schouder	0.80	8	4.93	44511	2.91	BS
305	shovel01.jpg	schep	1.00	5	4.71	30499	2.31	BS
306	sink.jpg	wasbak	0.67	6	4.81	36738	1.72	BS
307	skateboard.jpg	skateboard	1.00	10	4.52	31162	1.91	BS
311	smokingpipe.jpg	pijp	0.87	4	4.10	34145	2.78	BS
313	snowman.jpg	sneeuwpop	1.00	9	4.62	44235	1.68	BS
315	sock01a.jpg	sok	0.93	3	4.50	48826	2.14	BS
317	spatula03.jpg	spatel	0.93	6	4.40	32073	1.32	BS

318	spiderweb.jpg	spinnenweb	0.92	10	4.74	74968	1.34	BS
320	spoon01.jpg	lepel	1.00	5	4.60	30733	2.34	BS
321	springroll.jpg	loempia	1.00	7	4.59	43289	1.40	BS
322	squirrel.jpg	eekhoorn	0.92	8	4.69	49779	2.42	BS
324	stapler03a.jpg	nietmachine	1.00	11	4.50	33614	1.53	BS
325	starfish01.jpg	zeester	1.00	7	3.60	69191	1.23	BS
326	statue.jpg	standbeeld	0.93	10	3.67	48120	2.31	BS
327	steeringwheel.jpg	stuur	0.93	5	4.76	53253	3.70	BS
328	stool01.jpg	kruk	1.00	4	4.67	47507	2.05	BS
329	straw.jpg	rietje	1.00	6	3.70	37879	2.06	BS
330	strawberry.jpg	aardbei	1.00	7	4.60	59335	1.84	BS
331	suitcase.jpg	koffer	0.93	6	4.20	51557	3.17	BS
332	suitofarmor.jpg	harnas	0.83	6	3.98	85964	2.33	BS
334	surfboard.jpg	surfplank	0.80	9	3.61	38114	1.57	BS
336	swing.jpg	schommel	1.00	8	4.61	31500	1.85	BS
338	syringe02.jpg	spuit	0.87	5	3.51	45628	2.63	BS
339	table01.jpg	tafel	1.00	5	4.79	41901	3.56	BS
340	tank.jpg	tank	0.73	4	4.31	44011	2.93	BS
341	teabag.jpg	theezakje	0.87	9	4.40	36577	0.95	BS
342	tent.jpg	tent	1.00	4	4.45	51976	3.25	BS
343	thermometer02b.jpg	thermometer	0.87	11	4.10	37771	1.76	BS
344	tie02.jpg	stropdas	0.87	8	4.64	36063	2.29	BS
345	tiger02.jpg	tijger	0.93	6	4.36	52159	2.71	BS
346	tire.jpg	band	0.67	4	4.62	50392	3.54	BS
347	toaster01.jpg	broodrooster	0.93	12	4.50	44457	2.03	BS
348	tomato01.jpg	tomaat	1.00	6	4.70	45211	2.12	BS
349	tombstone.jpg	grafsteen	0.92	9	4.43	44915	1.96	BS
350	toothbrush03b.jpg	tandenborstel	1.00	13	4.70	40434	2.26	BS

351	tortoise01.jpg	schildpad	0.94	9	4.26	67239	2.28	BS
354	trampoline.jpg	trampoline	1.00	10	4.40	40355	1.53	BS
355	tray.jpg	dienblad	0.93	8	3.40	57572	1.74	BS
356	treadmill.jpg	loopband	0.93	8	4.67	51846	1.20	BS
357	tree.jpg	boom	0.87	4	4.69	75722	3.36	BS
359	tulip02.jpg	tulp	0.80	4	4.40	39409	1.34	BS
360	tweezers02a.jpg	pincet	0.93	6	4.30	31806	1.60	BS
361	umbrella04.jpg	paraplu	1.00	7	4.50	35263	2.18	BS
362	uprightpiano01.jpg	piano	1.00	5	4.67	58372	2.79	BS
363	usbkey.jpg	usb-stick	0.80	9	4.20	49500	0.85	BS
364	vacuumcleaner01.jpg	stofzuiger	1.00	10	4.50	31138	2.03	BS
365	vase01.jpg	vaas	1.00	4	3.70	34188	2.30	BS
366	violin.jpg	viool	0.93	5	3.60	45816	2.28	BS
367	wallclock.jpg	klok	1.00	4	4.67	59570	3.02	BS
369	walnut01c.jpg	walnoot	0.87	7	4.20	60575	1.51	BS
370	watch02a.jpg	horloge	1.00	7	4.40	44440	3.09	BS
371	waterfall.jpg	waterval	0.97	8	4.14	78885	2.25	BS
372	wateringcan.jpg	gieter	1.00	6	3.90	40685	1.18	BS
373	weight01.jpg	gewicht	0.67	7	3.80	44254	2.88	BS
374	wheelbarrow01.jpg	kruiwagen	0.94	9	4.41	47337	1.75	BS
375	wheelchair.jpg	rolstoel	1.00	8	4.52	66000	2.56	BS
376	windmill.jpg	molen	0.80	5	4.33	57445	2.30	BS
377	windshieldwiper02.jpg	ruitenwisser	0.86	12	4.32	35196	1.30	BS
379	woodboard.jpg	plank	0.93	5	3.80	37252	2.69	BS
380	woodenshoe.jpg	klomp	1.00	5	3.03	38515	1.60	BS
381	worldmap.jpg	wereldkaart	0.80	11	4.76	69769	0.78	BS
383	zebra.jpg	zebra	0.93	5	4.40	51331	2.13	BS
384	zipper.jpg	rits	1.00	4	4.83	38486	2.28	BS

09	Apenguin2.jpg	pinguïn	0.93	7	4.00	44758	2.03	B08
145	gingerroot.jpg	gember	0.73	6	4.00	56443	1.53	B08
310	slide.jpg	glijbaan	0.87	8	4.50	51116	1.36	B08
312	snail.jpg	slak	1.00	4	3.88	35637	2.02	B08
22	backpack_e2_s1.jpg	rugzak	0.80	6	4.38	31295	2.52	B13
220	mixer_e1_s1.jpg	mixer	0.67	5	3.38	28413	1.65	B13
230	newspapers.jpg	kranten	1.00	7	4.00	61695	2.96	MP
27	barn_owl.jpg	uil	1.00	3	3.75	43726	2.14	MM
30	bat.jpg	vleermuis	1.00	9	3.50	33901	2.15	MM
31	bathrobe.jpg	badjas	0.93	6	4.00	38651	2.00	MM
41	bone.jpg	bot	1.00	3	3.38	22348	2.80	MM
50	brain.jpg	hersenen	0.73	8	3.88	49450	3.19	MM
67	cap.jpg	pet	0.93	3	3.38	30980	2.76	MM
72	cauliflower.jpg	bloemkool	1.00	9	4.50	60186	1.40	MM
75	chess.jpg	schaakbord	0.67	10	4.13	51441	1.38	MM
77	church.jpg	kerk	0.87	4	3.50	59732	3.54	MM
82	coat.jpg	jas	0.67	3	3.00	38194	3.32	MM
109	doll.jpg	pop	0.93	3	3.00	43973	3.02	MM
149	glove.jpg	handschoen	1.00	10	3.63	43184	2.47	MM
152	grapes.jpg	druiven	0.87	7	4.75	56937	2.26	MM
165	harp.jpg	harp	1.00	4	3.38	43650	1.92	MM
192	lamp.jpg	lamp	0.93	4	3.75	33449	2.78	MM
231	nose.jpg	neus	1.00	4	4.88	41319	3.49	MM
250	pendants.jpg	oorbellen	1.00	9	3.25	38671	2.34	MM
261	plane.jpg	vliegtuig	1.00	9	4.63	28071	3.59	MM
263	pot.jpg	pan	0.80	3	4.50	45401	2.61	MM
267	pyramid.jpg	piramide	0.87	8	1.88	39139	2.09	MM
268	racket.jpg	tennisracket	0.80	12	4.38	46046	1.00	MM

308	skirt.jpg	rok	1.00	3	3.00	36567	2.50	MM
316	sofa.jpg	bank	1.00	4	4.75	45236	3.60	MM
333	sunflower.jpg	zonnebloem	1.00	10	4.38	59612	1.34	MM
337	sword.jpg	zwaard	0.87	6	3.38	28644	3.21	MM
358	trousers.jpg	broek	1.00	5	4.50	35613	3.47	MM
33	beanie.jpg	muts	1.00	4	3.63	70333	2.29	PB
73	cheese.jpg	kaas	1.00	4	4.00	57423	3.00	PB
137	football.jpg	voetbal	0.93	7	4.88	42195	2.75	PB
11	applepie.jpg	appeltaart	0.87	10	4.00	87580	2.06	WC
17	ashtray.jpg	asbak	0.93	5	3.50	69122	2.13	WC
51	bread.jpg	brood	1.00	5	3.88	91837	3.17	WC
58	burger.jpg	hamburger	0.87	9	4.38	69160	2.58	WC
79	cigarette.jpg	sigaret	1.00	7	3.50	26529	3.09	WC
100	cupcake.jpg	cupcake	0.87	7	4.38	55546	1.08	WC
107	dog.jpg	hond	1.00	4	4.50	38899	3.87	WC
112	door.jpg	deur	1.00	4	3.88	41623	4.03	WC
116	dress.jpg	jurk	1.00	4	3.75	98455	3.39	WC
150	goat.jpg	geit	0.73	4	3.63	46530	2.55	WC
157	gum.jpg	kauwgom	0.93	7	2.50	40601	2.41	WC
159	hair_straightener.jpg	stijltang	0.87	9	3.00	33563	1.11	WC
206	magazine.jpg	tijdschrift	0.67	11	2.13	84438	2.63	WC
223	motorcycle.jpg	scooter	0.80	7	3.50	56825	2.33	WC
232	olives.jpg	olijven	1.00	7	3.13	38543	2.01	WC
238	painting.jpg	schilderij	1.00	10	2.75	69877	2.97	WC
309	sleeping_bag.jpg	slaapzak	1.00	8	3.50	63855	1.85	WC
314	soap.jpg	zeep	0.93	4	2.38	32062	2.79	WC
319	sponge.jpg	spons	1.00	5	3.50	91840	2.04	WC
323	stamp.jpg	postzegel	1.00	9	2.88	92248	1.87	WC

335	sushi.jpg	sushi	1.00	5	4.00	42182	2.22	WC
352	traffic_light.jpg	stoplicht	0.87	9	4.88	33409	1.86	WC
353	train.jpg	trein	1.00	5	5.00	38889	3.51	WC
368	wallet.jpg	portemonnee	1.00	11	4.63	47446	2.74	WC
378	wine.jpg	wijn	0.87	4	3.75	33601	3.42	WC
382	yogurt.jpg	yoghurt	0.87	7	2.75	55733	1.98	WC

5 | Communicative intentions influence memory for conversations

Abstract

Speakers remember their own speech better than their interlocutors', a finding often attributed to the generation effect (Slamecka & Graf, 1978) and the production effect (MacLeod et al., 2010). However, when participants engage in natural dialogue, this advantage for self-produced speech does not always obtain (e.g. Knutsen & Le Bigot, 2014). Here we explore one of the differences between natural and experimental dialogue that may explain this phenomenon: communicative intention. This is cast in terms of the need to gain information, as when we seek information from someone else, that person's speech is put in focus, which research shows leads to a memory benefit. The current study uses question-answer pairs to manipulate both language production (speaking vs. listening) and communicative intentions (seeking vs. giving information). This allows us to test whether people remember their own speech better than their interlocutors' speech and whether that memory benefit changes when their interlocutors' speech is in focus. At study, thirteen participant pairs asked and answered questions about the location of three objects on the screen. At test, an online Yes/No recognition memory test conducted a day later, participants saw the names of all the objects presented at study. Preliminary results show that self-produced speech was remembered better than other-produced speech. Importantly, language production interacted with communicative intention, such that information-seekers remembered the answers to their questions better than information-givers remembered the questions they had been asked. That is, people's memory for comprehended speech can be moderated by how important they consider that speech to be. This sheds light on the role that communicative intention has on memory for conversations.

Introduction

Although one of the main functions of language is to facilitate interaction, relatively little is known about how conversations are represented and remembered by the interlocutors. The present study uses a controlled psycholinguistic paradigm to assess the influence of participant role (speaking vs. listening) and of communicative intentions (asking vs. answering a question) on memory for dialogue content.

One robust finding from research in memory for conversational content is that it varies as a function of participant role (speaker vs. listener): Speakers remember what they said better than listeners remember what they heard. This speaker advantage holds across a variety of stimuli: it has been explored for individual words and pictures (Brown, Jones, & Davis, 1995; McKinley et al., 2017; Yoon et al., 2016), as well as for sentences (Jarvella & Collas, 1974; Miller, 1996) and the cues used to generate sentences (e.g. "My favourite movie is...", Fischer et al., 2015). The speaker advantage holds for recall tasks (Miller, 1996), but has mostly been tested using recognition memory tasks (Fischer et al., 2015; Jarvella & Collas, 1974; McKinley et al., 2017; Yoon et al., 2016). It can even be found one week after study (Brown et al., 1995). Therefore, substantial evidence points to a speaker advantage for memory of conversational content.

The superior memory of speakers compared to listeners has been attributed to effects that originate in the mechanisms of language production, specifically the generation and production effects (Fischer et al., 2015; Hoedemaker et al., 2017; Knutsen & Le Bigot, 2014; McKinley et al., 2017; Yoon et al., 2016). The generation effect is the finding that coming up with a word provides a memory benefit relative to reading a word (Bertsch et al., 2007; Slamecka & Graf, 1978) and also relative to hearing a word (e.g., Dew

& Mulligan, 2008, Experiment 1A). This latter finding especially makes the generation effect applicable to conversation, showing that speakers have a memory benefit over listeners. The generation effect has been explained as increasing item-specific processing (Hunt & McDaniel, 1993) and making the memory trace of the item more distinctive (Gardiner & Hampton, 1988). The implication is that when people generate utterances, the generation effect allows them to remember their own contributions better than their interlocutors' utterances later on.

The production effect is the finding that saying words aloud improves memory relative to saying them silently, i.e., in inner speech, or not at all (MacLeod et al., 2010). This effect has also been attributed to distinctiveness, such that overt production provides additional information that a word had been said aloud (Ozubko et al., 2014). Therefore, because people speak during conversation, the production effect allows them to remember their own contributions better than their interlocutors' later on.

However, the dissociation between speakers and listeners in terms of content memory is called into question by more naturalistic studies, in which the generation effect and the production effect are reduced or even reversed (Hjelmquist, 1984; Knutsen & Le Bigot, 2014; Stafford & Daly, 1984, though see Miller, 1996). In one such study, Knutsen and Le Bigot (2014) asked participants to come up with a route that crossed certain points, which were marked on a map. Participants had 20 minutes to complete the task, after which they were instructed to write down as much of the conversation as they could recall. The authors reported more reuse of self-introduced referents (e.g., landmarks and street names) during the conversation, but no memory advantage for self- as opposed to other-introduced referents in the memory task. Hjelmquist (1984) gave participant pairs a topic (e.g., recent political events) and let them discuss it freely for five minutes. When participants were presented with sentences from this conversation four days later, they were equally good at recognising their own and their interlocutors' sentences. Similarly, Stafford and Daly (1984) had participant pairs get to know each other and then write down as much as they could remember. Here, participants recalled more information about their interlocutors than about themselves, reversing the typical speaker memory advantage.

It seems, then, that more tightly controlled studies report a speaker benefit, whereas more naturalistic studies do not. There are many differences between these two sets of studies that can affect memory and perhaps explain the discrepancy regarding the speaker benefit. For one, experimental conversations tend to have a speaker and a listener, unlike natural conversations, which tend to have two speakers. This is important for two reasons. Firstly, because conversations generally revolve around a topic, both speakers will inevitably reuse some of the same words (the ones integral to the topic at hand). That people reuse labels to refer to things is well-established: A label for an item needs to be introduced and negotiated to be accepted into common ground, but once that happens that word is reused even if it becomes over-informative (Brennan & Clark, 1996; Gann & Barr, 2014; Van Der Wege, 2009). Once a word is said by both interlocutors, they should both receive the memory benefit from having produced it. Indeed, Knutsen and Le Bigot (2014) found that higher numbers of reuses of a word was predictive of superior memory performance later on. Secondly, listening is considered more passive than speaking, meaning that the listener in an experimental conversation may not be very engaged in the conversation/task which could have detrimental effects on memory.

Experimental and natural conversations also differ in terms of discussion topics. Experimental conversations tend to be somewhat disjointed, e.g. discussing the location of objects in a grid (e.g. McKinley et al., 2017). Natural conversations, on the contrary, tend to be more cohesive, e.g. coming up with a route using a map (Knutsen & Le Bigot, 2014). This difference matters because, when there is a topic to clarify the relationship between sentences, information is integrated more easily, which aids later recall (Bransford & Johnson, 1972; Kozminsky, 1977). Additionally, when people are free to select a topic, they are likely to select a topic they find interesting or important, which can also improve memory (Brown et al., 1995; Keenan, MacWhinney, & Mayhew, 1977; Stafford & Daly, 1984). The lack of coherence could lead to overall poorer memory in experimental conversations, and heighten differences between listeners and speakers that would not exist in natural conversations.

Another major difference between experimental and natural conversations, and the main focus of this study, is the communicative intention associated with these different types of conversations. In most studies, participants' intention is to give information, like the order of pictures in a grid, or instructions *to* their interlocutors (McKinley et al., 2017; Yoon et al., 2016). In some more naturalistic studies, participants' intention is to get information *from* their interlocutors (Stafford & Daly, 1984). Arguably, when the participant's intention is to share information they themself know, like in the former scenario, emphasis is placed on the self-produced information. When the participant's intention is to get information from their interlocutor, like in the latter scenario, emphasis is placed on the other-produced information. Importantly, emphasised, or focused, information is known to be better remembered than non-focused information (Birch & Garnsey, 1995; Fraundorf et al., 2010; Sturt et al., 2004). This has been attributed to focused information receiving more attention or having a stronger or more accessible memory trace (Foraker & McElree, 2007).

Information can be put into focus in a number of ways: Most commonly, focus is expressed using intonation, but it can also be expressed using clefting and pseudo-clefting, focus particles (e.g. "only"), or by using context. In the current study, we manipulate focus by using question-answer pairs, as a word that answers a question is in focus and has been found to receive a memory benefit. Cutler and Fodor (1979) used a probe verification task in which they found that sounds in focused items were easier to detect than in non-focused items. Sturt et al. (2004) and Ward and Sturt (2007) used a change detection paradigm to test whether participants were better at detecting word changes when a word was focused compared to when it was not focused. This was indeed the case, but only when the focused word had been changed to a semantically related word (e.g. "beer" changing to "cider"). In Chapter 4 of this thesis, answers were found to be remembered better than the questions that elicited them were. We chose to use question-answer pairs because the way that they place focus on a referent is interesting in terms of the role of production in memory for conversations. This is a plausible case where the answer should be remembered very well by both interlocutors, for different reasons. The person asking the question (information-seeker) will receive a memory benefit as the answer is in focus and the person answering the question (information-giver) will receive a benefit because the answer was both generated and produced. Question-answer pairs are also very common (Graesser et al., 1994), making a good test case that bridges the results found in experimental and natural conversations.

As was mentioned previously, the beneficial effects of focus are often interpreted as a result of increased attention. Most of the research on this topic has tested how eye movements during reading differ between focused and non-focused items. The results of this literature are mixed, with some research reporting slowing-down when reading focused items, and some researchers reporting speeding-up when reading focused information (Birch & Rayner, 2010; Ward & Sturt, 2007). Benatar and Clifton (2014) argued that these differences are due to methodology and reported longer processing time for focused items. In visual world experiments examining speech planning, focused items have also been associated with increased look proportions (Ganushchak, Konopka, & Chen, 2014). That is, there is some evidence that focused items receive increased visual attention.

An interesting secondary function of focus is that it can trigger the computation of alternatives. In the case of question-answer pairs, this means that, for each question, there are different potential answers. Importantly, to the information-seeker, many potential answers can be equally plausible until an answer is given, whereas, for the information-giver, only one of the answers will (eventually) be suitable. For example, think of someone buying ice-cream. The ice-cream seller asking "Which flavour would you like?" may activate a couple of popular choices like"strawberry" and "chocolate", while the person buying the ice-cream may only briefly consider strawberry before settling on their favourite ice-cream flavour, chocolate. Therefore, by the time the answer "chocolate" is given, the ice-cream seller may have activated the concept of "strawberry" for longer. In this study, we explore the possibility that conversational role affects how alternatives are remembered after the conversation.

Current study

In the current study, we tested how speech production and communicative intention, as manipulated by question-answer pairs, affect memory for conversation. Our primary aim was to establish the extent to which pragmatic aspects of a conversation can influence memory, shedding light on some of the conflicting evidence regarding what speakers and listeners remember from conversations. Our secondary aim of this study was to test whether information-seekers remember alternative answers better than information-givers do. We also explored the link between focus and visual attention and its link to memory, with the goal of examining whether attention and focus play the same role in improving memory performance.

This experiment used a study-test paradigm. The study phase was performed in pairs with participants taking turns asking (information-seeker role) and answering (information-giver role) questions about the location of objects on the screen. Question-answer pairs were chosen because they not only manipulate communicative intention, but they also allow the engagement of two participants in an interactive task, rather than having one participant be the speaker in one trial and the other participant be the speaker in the next. Participants could see the same three objects on the screen, of which one was used in the question, one in the answer, and one was not mentioned. An example of a trial can be seen in Figure 5.1. In this example, Speaker A, who is the information-seeker, asks “What should go next to the apron?” and Speaker B, who is the information-giver, answers “The accordion”. Note that, for the information-seeker, the unmentioned item (the accordion) would be a plausible alternative to the target, whereas for the information-giver it would not, as this speaker already knew the correct answer. Both participants’ eyes were tracked in the study phase. The test phase was an online Yes/No recognition memory task in which participants saw the names of each picture presented in the study phase intermixed with an equal number of new words.

We predicted that participants would remember the words they said themselves better than the words their partners said. In the example in Figure 5.1, Speaker A would remember “apron" better than “accordion". We

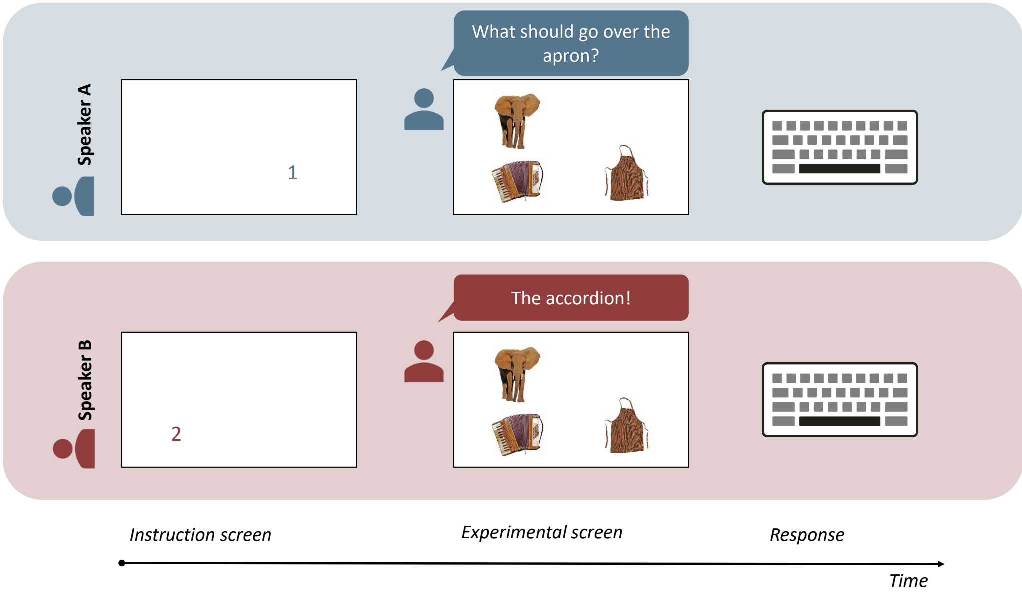


Figure 5.1: Example of a trial in the study phase.

also predicted that information-seekers would remember the words produced by their partner better than the information-givers would. That is, Speaker A would remember “accordion" better than Speaker B would remember “apron". Finally, we predicted unmentioned items would be remembered worse than mentioned items. That is, “elephant” would be remembered worse than “accordion” and “apron” by both Speaker A and Speaker B.

We also planned two exploratory analyses. More specifically, we planned to explore the possibility that information-seekers would remember unmentioned items better than information-givers did. We also planned to explore our eye-tracking data to investigate further the link between visual attention and memory.

Methods

Participants

We aimed to recruit 96 individuals in 48 pairs to participate in this experiment. This sample size was determined by running a power analysis, in which we simulated data using an effect size of approximately 20% for the combined effect of production and generation and of 6% for the effect of focus, following the results reported in Chapter 4. With those parameters, we would have 80% power to observe a significant interaction, showing that information-seekers remember answers better than information-givers remember questions. Due to the COVID-19 outbreak, data collection was terminated early after testing 26 individuals in 13 pairs. As such, the analyses reported here are likely underpowered.

Participants were recruited from the Max Planck Institute for Psycholinguistics participant database and compensated 10 €. All participants (23 female) were 18-30 years old (M = 23) and were native Dutch speakers with no reported speech or language problems and with normal or corrected-to-normal vision. Ethical approval to conduct this study was given by the Ethics Board of the Social Sciences Faculty of the Radboud University.

Materials

Stimuli were the 384 colour photographs used in Chapter 4, which were normed for name agreement, familiarity, visual complexity (measured in JPEG size), $\log_{10}$ frequency, and length (measured in letters). Items were first split into two matched sets (Sets A and B), each of which was split further into three matched subsets (Subsets A1, A2, A3, B1, B2, and B3). A full description of the norming and matching procedure can be found in Chapter

4. The matched Sets and Subsets were used to counterbalance the allocation of items to probe type (target or foil) and item type (answer, question, or unmentioned) conditions in six lists.

In the study phase, participants saw 192 photographs, of which 2/3 were named in the process of asking and answering questions and the remaining 1/3 were not mentioned. In the test phase, the names of the 192 photographs shown at study (targets) were shown intermixed with an equal number of new words (foils) for a total of 384 words.

Apparatus and procedure

Study phase. The study phase was conducted in an experiment room at the Max Planck Institute for Psycholinguistics. Participants were sitting side by side 55 cm in front of their own monitor (1920 x 1080 resolution). They were able to hear but not see each other. The experiment was controlled using Presentation (version 18.3; Neurobehavioral Systems, Berkeley, CA, USA). Vocal responses from both participants were recorded using a microphone. Eye data were collected using an EyeLink 1000 Plus (SR Research Ltd., Osgoode, Canada) eye-tracker.

At the beginning of the session, participants gave informed consent and then completed random-order 9-point calibration and validation routines. They then completed eight practice trials followed by 64 experimental trials. Participants received feedback during the practice trials, which were otherwise identical to the experimental trials. Participants had the option of a short break after 32 experimental trials, which none of them took.

Each trial started with an instruction screen, which consisted of a blue "1" for the information-seeker and a red "2" for the information-giver. The position in which those numbers appeared signalled what items speakers should use in their utterances. In the example in Figure 5.1, "apron" was used in the question because it appeared where the “1" was previously. The same applied for the answer. After 2500 ms, the instruction screen was replaced by the experimental screen, consisting of the three images, which appeared in the same positions for both speakers. The information-seeker then had to find the empty space (to know what preposition to use) and asked the question, in this case "What should go over the apron?". The information-giver would then give the answer, in this case "The accordion". Finally, both participants pressed the space bar. The following trial started when the experimenter pressed a button on a button box. This was done to avoid the two systems going out of synchronisation. Trial order was manually randomised.

Test phase. The test phase was a self-paced Yes/No recognition memory task conducted online 20-28 hours after the study phase[1]. In this phase, participants read the names of each of the images they had seen in the study phase intermixed with an equal number of new concrete nouns. Their task was to press “Yes" when they saw words referring to pictures they had seen in the study phase, including unmentioned items, and to press “No" when they saw words referring to items they did not remember seeing in the study phase.

Analysis

The dependent variable in all analyses was memory performance (“Yes”: 1; “No”: 0). Memory performance was analysed using logistic mixed-effects regression models run in R (version 3.6.1; R Core Team, 2019) and implemented in the lme4 package (version 1.1-21; Bates et al., 2015) using the optimiser *BOBYQA* (Powell, 2009). There were two confirmatory analyses and

[1]Two participants completed the test phase a little later, 29 and 30 hours after the study phase.

two exploratory analyses, which were preregistered on the Open Science Framework (https://osf.io/y7seu/registrations).

Predictors in the main confirmatory analysis were the participant's communicative intention (information-seeker vs. information-giver) and item condition (self-produced, other-produced, unmentioned). The intention predictor assessed the effect of communicative intention (to gain or to give information) and the item predictor tested the effect of generation and production (self- vs. other-produced) and of mention (mentioned vs. unmentioned). This analysis was run on the targets only, as foils did not belong to any of these conditions. A separate confirmatory analysis examined the effect of probe type (target vs. foil) to ensure that participants were able to distinguish between old and new items.

An exploratory follow-up analysis was run on the unmentioned items only with intention as the predictor, to test whether information-seekers were more likely to consider these items as alternatives. An additional exploratory analysis was run on all targets to assess the influence of visual attention, as measured by gaze duration, on memory. This analysis used all the predictors from the main analysis plus the continuous predictor gaze duration.

Trials were excluded when a picture was named incorrectly by either participant in the pair. In the cases when the incorrect name appeared elsewhere in the experiment, both instances of the word were excluded. This led to the exclusion of 129 trials (out of 1664; 8% of the data). Other preregistered exclusion criteria included participants not completing the second phase of the experiment, but that never occurred.

Results

Memory performance overall was 64% (*SD*: 48%). Consistent with the findings reported in Chapter 4, answers were remembered better than questions, which were in turn remembered better than unmentioned items. More specifically, words that appeared as answers were recognised correctly 56% (*SD*: 50%) of the time, words that appeared as questions 47% (*SD*: *SD*: 48%) of the time, and words that appeared as unmentioned items 23% (*SD*: 42%) of the time. In the present study, we were more interested in how speakers remembered different items depending on their role in that conversation, so we did not check if this difference is statistically significant. To preview our results, the generation effect and the production effect continued to play a role in the present study, with self-produced items being remembered the best. We also found evidence that communicative intentions and focus play a role, such that participants asking a question remembered the answer better than participants giving an answer remembered the question. A visualisation of the results can be seen in Figure 5.2

The first confirmatory analysis tested the effect of probe type (targets = 0.5, foils = -0.5) to ensure that participants could distinguish between old and new items. The random structure included by-participant and by-item intercepts and by-participant and by-item slopes for the targets vs. foils contrast. The results of this analysis can be seen in Table 5.1. The negative intercept suggests that participants had a negative response bias, i.e., they were overall more likely to respond "No". The positive estimate for the targets vs. foils contrast suggests that participants were more likely to respond with "Yes" to targets than to foils, i.e., were able to distinguish between old and new items.

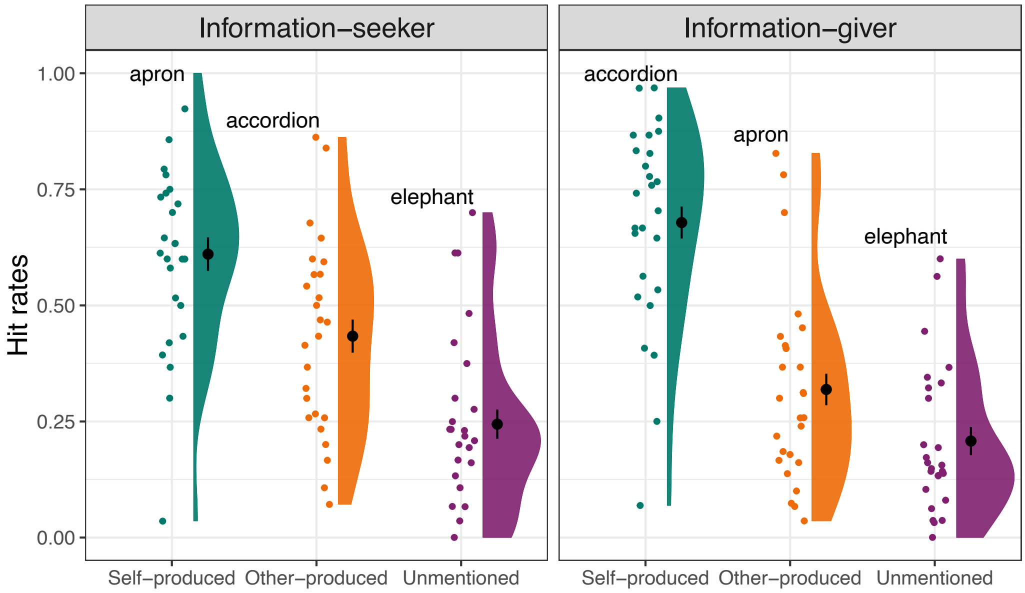


Figure 5.2: Hit rates in the memory task to each Item condition split by the Communicative intention of the speaker. The annotations reference the example in Figure 5.1, in which the information-seeker asked "What should go over the apron", and the information-seeker responded "The accordion". The information giver (in the right panel) remembers the accordion better than the apron. The information-seeker (in the left panel) remembers the apron better than the accordion. That is, both speakers benefit from the generation effect and the production effect. However, the difference in hit rates for the apron and the accordion is smaller in the case of the information-seeker, because focus makes the accordion more memorable. Each coloured point represents a participant's mean hit rate for that condition. The black points represent the overall mean hit rate for that condition. The bars around the black point represent the normalised within-participant 95% confidence interval.

The second confirmatory analysis tested how communicative intention and item condition can affect memory in conversations. Intention was deviation contrast-coded (information-seeker = 0.5, information-giver = -0.5). Item-condition was Helmert contrast-coded and split into two contrasts: one testing the effect of having been mentioned (by self or other = 0.25, by no one = -0.5) and one testing production (self = -0.5, other = 0.5). The random structure for this model included by-subject and by-item intercepts only. Including random slopes led to singular fits, presumably due to the low number of participants in this study. The results of this analysis can be seen in Table 5.2. A visualisation of the underlying data can be seen

Table 5.1: Mixed-effects logistic regression testing the effect of Probe, i.e., targets vs. foils.

Fixed effects					
	Estimate	SE	Wald z	p	*CI*
Intercept	-1.10	0.14	−8.05	< 0.001	−1.37, −0.82
Targets vs. Foils	1.39	0.11	12.34	< 0.001	1.16, 1.62
Random effects					
		Variance	SD		
Participant	Intercept	0.45	0.67		
	Targets vs. Foils	0.23	0.48		
Item	Intercept	0.10	0.31		
	Targets vs. Foils	0.26	0.51		

in Figure 5.2. The negative intercept shows a slight "No" response bias. This is a little surprising, considering this analysis was run exclusively on targets, i.e. old items. However, targets include unmentioned items, which were associated with very low recognition rates. The positive estimate for communicative intention shows that information-givers had an overall memory advantage for what was said in a dialogue. The positive estimate for the mention contrast shows that, unsurprisingly, mentioned items were remembered much better than unmentioned items. Additionally, the estimate for the production contrast shows that participants were better at recognising items they had come up with and said themselves. The estimate for this contrast is negative because the contrast for self-produced items was -0.5. That is, this finding is evidence that speakers benefited from the generation effect and the production effect. Importantly, the interaction between communicative intention and production shows that there is a difference between how well other-produced words are remembered by information-seekers and information-givers. More specifically, it seems that information-seekers (who asked the questions) remembered the answers better than information-givers (who gave the answers) remembered the questions. This finding provides evidence for the effect of focus.

Table 5.2: Mixed-effects logistic regression testing the effect of Communicative intention, i.e., information-seeker vs. information-giver and Item condition. The Item condition predictor was split into Mention (mentioned vs. unmentioned) and Production (self-produced vs. other-produced).

Fixed effects					
	Estimate	SE	Wald *z*	*p*	*CI*
Intercept	−0.47	0.17	−2.74	0.01	−0.82, −0.12
Information-seeker vs. Information-giver	0.16	0.07	2.32	0.021	0.03, 0.30
Mentioned vs. Unmentioned	1.99	0.11	18.68	< 0.001	1.78, 2.20
Self-produced vs. Other-produced	−1.32	0.08	−15.69	< 0.001	−1.49, −1.16
Info-seeker vs. giver:Ment. vs. Unment.	−0.16	0.21	−0.78	0.437	−0.57, 0.25
Info-seeker vs. giver:Self- vs. Other-prod.	0.93	0.17	5.60	< 0.001	0.60, 1.25

Random effects			
		Variance	SD
Participant	Intercept	0.71	0.84
Item	Intercept	0.23	0.48

The interaction between communicative intention and mention tested whether there was a difference in how unmentioned items were remembered by information-seekers as opposed to information-givers. There was no evidence supporting that that information-seekers remembered unmentioned items better than information givers did ($M_{information\text{-}seeker}$ = 0.24, $SD_{information\text{-}seeker}$ = 0.43; $M_{information\text{-}giver}$ = 0.21, $SD_{information\text{-}giver}$ = 0.41).

A further exploratory analysis looked at the relationship between visual attention at study, measured by gaze duration, and memory performance at test. This analysis was based on the idea that increased attention is behind the processing benefits associated with focus. This was explored in a further model that included gaze duration as a covariate. Gaze duration was centred and scaled to be entered in the model and the rest of the predictors were contrast-coded as described previously. There was no evidence that gaze duration was a good predictor of memory ($\beta = 0.06$, $SE = 0.04$, $z = 1.53$, 95% CIs [-0.01, 0.15], = $p = 0.11$). Numerically, it does seem like focused items received longer gaze durations (M_{answer} = 2445 ms, SD_{answer} = 5830 ms; $M_{question}$ = 1902 ms, $SD_{question}$ = 3163 ms; $M_{unmentioned}$ = 785 ms, $SD_{unmentioned}$ = 2074 ms).

Discussion

In this study, we investigated two factors that can influence memory for conversations: 1) speech production and 2) focus, as a result of communicative intention. We tested the effects of these two factors by having participants take turns asking and answering questions about the location of objects on a grid and then testing their memory for the names of those objects. We predicted that self-produced speech would be remembered better than other-produced speech, but that this difference would vary depending on the speaker's communicative intention (information-seeker vs. information-giver). More specifically, we predicted that information-seekers would remember other-produced speech better than information-givers would, because conversational goal puts other-produced speech (the answer) in focus. Data collection could not be completed due to the COVID-19 outbreak, but preliminary results support our hypotheses: Participants generally remembered self-produced speech better than other-produced speech, but this difference was smaller for information-seekers. We explain the benefit for self-produced speech as a combination of the generation effect and the production effect. We explain the benefit for other-produced speech by information-seekers as an effect of their communicative intentions placing focus on other-produced speech.

A clear finding from the present study is the benefit for self-produced speech over other-produced speech. This finding can be readily explained by the generation effect and the production effect, which have been successful at predicting such effects in one-person studies and in recent studies of dialogue (e.g., Fischer et al., 2015; McKinley et al., 2017; Yoon et al., 2016). Both the generation and the production effect can be thought of as increasing item-specific processing, making the traces of self-produced words better remembered than other-produced speech (Begg et al., 1989; MacLeod et al., 2010; Ozubko & MacLeod, 2010). We found that the benefit for self-produced speech remained quite large even when taking into account factors benefiting other-produced speech.

The main finding of this study is that memory for other-produced speech reliably improved when it was placed in focus. Focused items are thought to receive more attention, which leads to improved memory, among other benefits (Foraker & McElree, 2007; Sturt et al., 2004). Here, we manipulated focus through communicative intentions, i.e., we hypothesised that speech would be considered important when it fulfilled a speaker's goal in a conversation. Participants' intention in this experiment was to gain information by asking questions, which put the answers in focus. As such, the participants are assumed to have been paying close attention to these answers, which improved their memory for them relative to comprehended speech that was not focused. By manipulating focus, we were able to shrink the benefit typically observed for self-produced over other-produced speech.

Following the idea that questions put answers in focus, we explored the possibility that alternative answers (i.e., unmentioned items) would be remembered better by information-seekers than by information-givers. This was not found to be the case. There are several reasons as to why that might be. Perhaps the information-seekers did not really consider these items as alternatives, or, if they did, they suppressed them once the correct answer was given. It is also possible that we do not have enough power to observe this (probably small) effect, as we were not able to complete data collection. Collecting more data, once possible, may help distinguish between these possibilities.

In a second exploratory analysis, we tested whether visual attention during processing predicted superior memory at test and again found no significant effect. As before, the reason behind this null finding may be low power, given that we ended data collection prematurely. It is also possible that gaze duration during the entirety of the trial is not a sufficiently sensitive measure of visual attention for language processing. In Chapter 3 of this thesis, we found a close relationship between memory and processing time (as measured by naming latencies), but gaze duration during the entirety of the trial did not explain additional variance. Again, collecting more data, once possible, may help distinguish between these possibilities.

While it was a step towards examining the relationship between memory and language during conversation, this study did have some limitations. Firstly, the communicative goal used was artificial—participants were not intrinsically motivated to find out how the different items should be arranged and the answers did not serve a function later in the experiment. Studying more naturalistic communicative goals and intentions would not only bring us closer to the reality of everyday communication, but it could also have implications for the size of the effect. That is, when intentions are intrinsic they might have an even larger effect on memory, further closing the gap between memory for self- and other-produced speech. Future research could address this problem by allowing participants to pick the answer themselves or by making those answers relevant for the upcoming trial.

Despite these limitations, this work makes an important contribution to research on dialogue. Firstly, we examined focus in a novel way, by embedding it in a conversation and tying it to communicative intentions. That is, we assumed that people can use focus to flag what is important to them in a conversation and found that intentions can affect what people remember from their conversations. This shows that the reasons why people produce utterances in conversations can influence how well the words within the utterances are remembered.

The finding that communicative intentions and focus influence memory for conversations also highlights the importance of studying the 'listeners' in a conversation. Unlike natural conversations which generally involve two speakers, psycholinguistic studies of dialogue often involve a speaker, who is responsible for 'achieving' the goal of a trial, and a listener. As such, it is perhaps not surprising that speakers remember events during experiments better than listeners do, as the speakers are the ones that are most engaged during the experiment. However, to develop a better understanding of what affects memory during conversations, we need conversations that simulate the dynamics between speakers better. Using the concept of communicative intentions, we were able to make comprehended speech more or less important to the participants, modulating the recognition of other-produced speech. This work demonstrates the need for studies where both (or all) speakers in a conversation have an active role to play: not only does this better replicate natural conversation in the lab, but it has important implications for what is remembered during the experiment.

In sum, we have shown that speech production and communicative intentions both improve memory for words in a conversation. When a joint communicative goal is accomplished by a conversation partner, additional attention is allocated to their speech, such that memory for it improves. However, that improvement is not as large as the memory improvement stemming from coming up with and saying words aloud. This work makes important advances to the understanding of what people remember from their conversations.

6 Summary and discussion

The aim of this thesis was to investigate some of the factors that influence what people remember from their conversations, thus linking the fields of memory and language. In order to address that aim, this thesis focuses on two topics, the asymmetry between language production and language comprehension and the importance of studying language in a dialogue context.

The first topic is that language production and language comprehension, i.e., the linguistic components of dialogue, have different effects on long-term memory, such that produced language is remembered better than comprehended language. This asymmetry in memory retention is especially surprising given that recent psycholinguistic theories posit parity of representations between production and comprehension (e.g., Pickering & Garrod, 2004, 2013). To understand where this difference comes from, Chapters 2 and 3 investigated the processes of language production that enhance memory. Chapter 3 also examined the level of representation benefiting as a result of word generation. As discussed below, this work bridges research on psycholinguistics and on memory and has theoretical implications for both.

The second topic is that communication is usually not studied in psycholinguistic research, despite it being one of the main functions of language. The joint investigation of language and memory in a conversational context, the subject of this thesis, is especially important because learning often happens in such contexts. This thesis aimed to make theoretical and methodological contributions to the study of language in a naturalistic setting. As such, the studies described in Chapters 4 and 5 developed a paradigm that uses a linguistic structure that is frequent in everyday communication and that engages two speakers. Although the dialogue examined in this thesis is still fairly artificial, the present work illustrates how to introduce elements of natural conversation into experimental work. This work is relevant to many everyday situations, from getting insight into differences in personal recollections of conversations, to educational settings and word learning.

In the following sections, I will first summarise and evaluate the findings from each experimental study. I will then discuss the implications of this work for memory phenomena and for the joint study of language and memory in terms of word production processes and conversation. I also note potential avenues for future research.

Summary and evaluation of individual chapters

In Chapter 2, I tested whether both coming up with words and saying them aloud improve memory in a picture naming task. In this experiment, participants first named pictures silently or aloud with the picture names or unreadable labels superimposed and then performed a recognition memory task with the pictures they had named. In the memory task, participants performed better on the pictures they had named aloud, as opposed to pictures they had named silently, replicating the production effect. They also performed better on the pictures they had named themselves, as opposed to pictures the names of which they could read, replicating the generation effect. These findings bridge research on memory phenomena, on the one hand, and on conceptually mediated language production, on the other. More specifically, they show that the generation effect arises before or during lexical processing and the production effect arises in post-lexical processing. This finding is important for language research because it provides some explanation as to why people remember the words they say better than the words they hear.

The findings of Chapter 2 also have implications for memory phenomena. First, this chapter extends the boundaries of the generation effect to include conceptually mediated language production that does not rely on orthography or phonology to aid retrieval (see also Weldon & Roediger, 1987). Second, it reveals a potential confound for the picture superiority effect, or the finding that pictures are remembered better than words (Paivio et al., 1968). I showed that part of the memory benefit that is attributed to the picture superiority effect is due to the generation effect instead: Accessing a word form from a picture is different to accessing it from a word, as the former requires generation. This is especially relevant for the interpretation of the results of studies, like Fawcett et al. (2012), in which participants produce words from pictures vs. from their names.

In Chapter 3, I showed that the generation effect generalises to situations when participants recognise words instead of pictures. That is, participants were more accurate at recognising the names of pictures they had named themselves as opposed to the names of pictures the names of which they had heard and then repeated. This finding shows that the origin of the generation effect in picture naming is not purely visual: If only visual representations were enhanced during generation, a memory benefit should not have been observed when participants were tested on the names of the pictures. This finding is further discussed in the section 'Relationship between memory and word production processes'. That words receive a memory benefit is also important for the goal of understanding memory for conversations, as most conversations do not have distinct imagery (like a picture) associated with each element or episode. As such, testing people's memory for the words that were said is more consistent with the phenomena this thesis is interested in.

In Chapter 3, the speed with which the pictures were named, the naming latency, was a very good predictor of memory such that the slower a word was named, the better it was remembered. This was only the case when participants generated the picture names themselves: Slower naming when participants had heard the picture names and repeated them led to no memory advantage. That is, longer conceptual or lexical processing was beneficial to memory, but other types of processing were not. This finding demonstrates that conceptual or lexical processing plays an important role in the generation effect. Perhaps surprisingly, total gaze duration to pictures did not predict additional variation after effects of naming latency were accounted for. This suggests that visual processing of pictures, when not paired with conceptual or lexical processing, has minimal influence on recognition performance, at least recognition of picture names. This finding highlights the importance of taking into account different levels of processing in memory research. The involvement of word production processes in the generation and the production effect are further discussed in the section 'Relationship between memory and word production processes'.

An unexpected finding in Chapter 3 was the interaction between naming latency and the contrast comparing the neutral generation condition with the unrelated distractor condition, suggesting a complex role for lexical competition in memory. Although naming latency was associated with a memory advantage in both conditions, that advantage was smaller in the distractor condition than in the neutral generation condition. This went against the expectation that longer naming latency would lead to memory benefits in both conditions in the same way. To the extent that increased processing time in the distractor condition reflects increased competition between the distractor and the target, this implies that some competition led to a memory benefit, but as that competition increased, the increased effort led to diminishing returns.

In the second part of this thesis, I turned my attention from language production to communication. In Chapters 4 and 5, I used question-answer pairs to manipulate focus in a manner that is common in everyday communication. This way, I was able to test how well the focused, i.e. most important, elements of an utterance are remembered relative to non-focused elements. By manipulating whether a participant was asking or answering a question, I was also able to test how memory for the focused word can vary depending on the interlocutor's conversational role.

In the experiment presented in Chapter 4, participants saw three objects on the screen and listened to recorded question-answer pairs mentioning two of the presented items. In one such trial, participants would see the picture of a painting, a goat, and a doll and hear the interaction "What should move next to the painting?" "The goat." When, a day later, participants were tested on the names of the pictures they had seen, they showed better recognition for items mentioned in the answer than ones mentioned in the question. That is, questions were found to place focus on answers such that answers were remembered better than the questions themselves. This is noteworthy considering that participants had not actively contributed to the conversation, but passively (over)heard it. This finding highlights focus as a very powerful way of improving memory in comprehension. Until now, memory for focused information has been investigated in relation to the comprehension of sentences and longer, usually written, narratives. As such, the present work is novel in its use of focus to study dialogue.

A strength of the paradigm of the experiment detailed in Chapter 4 is its use of question-answer pairs, a structure that is common in everyday communication. This contrasts with most research on focus, which has typically used structures that are extremely rare in conversation, like clefts and pseudo-clefts. In addition, in this paradigm participants were tested on single words, the names of pictures presented in the study phase, e.g., painting, goat, doll. This is useful in that it makes it possible for all items from the study phase to be tested in the same way. However, a limitation is that it compresses the meaning of a sentence to a single word, thus losing some nuance. This paradigm may be easiest to use with simple sentences, like the ones used here, but could be extended to investigate larger discourses or more complex structures.

Chapter 5 used question-answer pairs like the ones in Chapter 4, with the difference that, in this case, one participant asked the question, and the other participant answered it. This allowed me to investigate the effect of language production (speaking vs. listening) and the effect of focus (asking vs. answering a question) on memory for conversations. Here, focus was tied to the notion of communicative intentions, captured by the roles of information-seeker and information-giver. The findings of Chapter 4 already showed that questions put the answers in focus; here I assume that people ask questions because of their communicative intention to gain information. Both language production and focus continued to influence memory during conversations. Participants remembered what they had said themselves far better than what they had heard their interlocutors say, replicating the findings of Chapters 2 and 3. However, the size of that benefit depended on the participant's role in the conversation and was smaller when participants were in the information-seeker role. That is because information-seekers remembered the answers, which they heard, better than information-givers remembered the questions.

The results of Chapter 5 show that focus can play an important role in language processing during dialogue. They also show that comprehension is not a passive process during which a listener merely receives instructions or directions from their interlocutor. On the contrary, listeners allocate more attention to, and as a result remember better, speech that they consider to be in focus as opposed to speech that they do not. These findings are discussed further in the section 'Memory for conversation'.

Memory phenomena

A large part of this thesis was concerned with linking language to memory phenomena, specifically the generation effect, the production effect, and, to a degree, the picture superiority effect. These memory phenomena, along with many others, share a number of similarities and, as a result, theories have often sought to explain them together.

One such account that has been applied to multiple phenomena, and which I have broadly followed here, is the distinctiveness account (Hunt & McDaniel, 1993; Hunt & Worthen, 2006). According to this account, when items are studied in an "unusual" way (e.g. generating vs. reading, producing aloud vs. silently), that mode of study is encoded and used heuristically at test. Although the work presented here was not designed to evaluate the distinctiveness account, the findings of Chapter 2 could have some implications for it. In this chapter, old and new pictures were presented in a recognition memory test as they had in the study phase, i.e. with labels superimposed (prompting participants to read or to generate words) and with red and green frames around them (prompting participants to say the words silently or aloud). If participants used the distinctive study modes heuristically at test, e.g. remembering “I came up with this word” or “I said this word aloud”, then they should have made fewer false alarm errors in these conditions than in the non-distinctive conditions (see also Dodson & Schacter, 2001). However, there was no evidence of this in the bias measures, which in fact suggest that participants had a generally positive response bias for distinctive conditions. As such, these results do not provide support for the hypothesis that distinctiveness can be used heuristically at test.

The work presented in Chapter 3 provides some evidence for theories suggesting that increased effort or *desirable difficulties* are beneficial to memory (Bjork, 1994). According to the *retrieval effort hypothesis*, effortful, but successful, word retrieval is associated with memory benefits (Pyc & Rawson, 2009). This account could easily explain the findings of Chapter 3, where words that were retrieved slowly were found to be remembered better than words that were retrieved quickly. Is it possible then that desirable difficulties play a role in the generation effect?

One way to look at this is by comparing the effort required by *different* processes: The generation effect may arise because reading words is easier than generating them. That seems unlikely, however, as generation tasks that are considered difficult (e.g. multiplication) or easy (e.g. letter switching) lead to similar memory benefits (Bertsch et al., 2007). The results of Chapter 3 also speak against this interpretation, as longer processing did not lead to a memory benefit when generation was not required. Another way to look at the role of desirable difficulties is by comparing the amount of effort required by different instances of the *same* process: The effort expended when generating a difficult word as opposed to an easy word may affect the size of the memory benefit. The findings of Chapter 3 provide tentative support for this hypothesis.

This interpretation of the results assumes that the generation effect arises during processes that occur when generating words but not when reading them and that these processes interact with effort or desirable difficulties. This is a potential extension of the explanations of the generation effect. It is possible that these generation processes are affected by effort in different ways: Conceptually mediated word generation, like in Chapter 3, may benefit from effort more than other kinds of generation, like multiplication. According to this view, effort is not the cause of the generation effect, but it can moderate it. This hypothesis could be tested by varying the difficulty of word retrieval within different generation tasks.

Relationship between memory and word production processes

The findings of Chapters 2 and 3 raise interesting questions regarding the relationship between memory and word production processes. The work presented in these chapters shows that word production processes can interface with episodic memory, but a question that remains is which conceptual or lexical processes give rise to the generation and production effects.

The findings of Chapter 3 demonstrate the importance of conceptual or lexical processes, as opposed to visual processes, for the generation effect but cannot distinguish further which processes are involved. Going back to the Levelt et al. (1999) model discussed in Chapter 1, some possibilities include conceptual preparation, lexical access, and form encoding. Evaluating the involvement of each of these processes is difficult, but some suggestions are listed below.

The influence of conceptual preparation could be tested with pictures that vary in name agreement. For example, a picture of an apple is almost always called “appel”, but “friet” or “pataat” are both equally good labels for a picture of fries. The prediction here is that words with multiple viable labels will be named slower; if that additional time spent naming is predictive of later (positive) memory performance, that would indicate that conceptual preparation plays a role in the observed memory benefit. Similarly, form encoding could be addressed by inducing tip-of-the-tongue states; if longer naming latencies are associated with better memory that would indicate a role for form encoding in the memory benefit.

The effect of lexical competition on memory might be underscored by a surprising finding from Chapter 3. This is the finding that longer naming latencies were associated with a memory benefit, but that benefit was larger for the neutral generation condition than the (more competitive) distractor generation condition. So what does this mean about lexical competition? As this interaction between naming latency and the comparison between the neutral generation and distractor conditions was not predicted, it should be replicated. If this finding obtains in a second study, it could be investigated further using words with sparse or dense semantic and phonological neighbourhoods, which are subject to different levels of competition.

Turning to the production effect, the findings of Chapter 2 suggest that this emerges in post-lexical processing. This is consistent with tests of the production effect, in which participants read or hear a word and then produce it in different manners (e.g. say it aloud, sing it, write it). Although the meaning of these words may be accessed, conceptual and semantic processing is not required to access the word-form. Similarly, the finding that the production effect is observed with non-words (MacLeod et al., 2010) strongly suggests that conceptual and semantic processing is not involved in the production effect.

From work in psycholinguistics, we know that after the form of the word has been accessed, it needs to be phonologically and phonetically processed and, eventually, articulated. There is some evidence that the volume at which a word is produced has an effect on the size of the effect (Quinlan & Taylor, 2013), suggesting that phonetic encoding (to the extent that different manners of speaking affect the phonetic realisation of a word) and articulation may be of more importance to the production effect than phonological encoding.

The findings discussed here constrain what the interface between word production processes and episodic memory might be. However, it is currently unclear what that interface looks like. Though models of distinctiveness exist (Jamieson, Mewhort, & Hockley, 2016), they generally do not take into account multiple levels of word representation, which would be needed to observe the pattern of results described in Chapters 2 and 3. This could be addressed in future work.

Memory for conversation

The central aim of this thesis was to get a better understanding of the factors affecting memory for speaking and for listening, with the goal of examining these factors in conversational settings. Chapters 2 and 3 showed that produced speech is remembered better than comprehended speech because two aspects of word production, generation and speaking aloud, have a beneficial effect on memory. Chapter 4 showed that not all parts of comprehended speech are remembered equally well—the focused parts of a conversation are remembered better than the non-focused ones. Finally, Chapter 5 integrated the previous findings to show that, although produced speech is remembered better than comprehended speech, the size of that benefit depends on a participant's role in the conversation. Combined, these findings show the importance of situating research on memory and language in conversational settings.

The main contribution of this thesis on the study of memory for dialogue is the finding that focus, guided by communicative intentions, can affect how well comprehended speech is remembered. This was shown in Chapter 5, where participants who asked a question (information-seekers), remembered the answer better than participants who answered a question (information-givers) remembered the question. The idea here is that the goal of information-seekers is to find out information, which makes them pay attention to the speech they are comprehending. This is important because it shows that comprehension entails more than passively receiving information and that people engaging in conversation evaluate which parts of the conversation are more interesting or relevant to them. As such, some of the reported asymmetries between language production and comprehension may be, partly, due to the emphasis that is given to speakers in most psycholinguistic paradigms.

This thesis also involved the development of a paradigm that facilitates the investigation of dialogue without relinquishing experimental control. Aspects of this paradigm were developed in Chapters 2 through 4 and combined in Chapter 5. The studies in Chapters 2 through 4 tested factors that affected how well individual participants remembered words they produced or heard. An assumption here is that the linguistic and memory processes assessed in these paradigms also play a role in interactive contexts. The study in Chapter 5 combined the previously studied processes and tested that they obtain in a pair study. This approach can allow relative flexibility when testing pairs by first testing and describing effects on individuals.

In the paradigm presented in Chapter 5, participants were tested on how well they remembered questions and answers they had produced or comprehended, situating a recognition memory test within a simple dialogue. A strength of this design is that it engaged two participants to a similar extent in each trial. This is important because most studies comparing memory for produced and comprehended speech have participants act as a speaker or a listener—the former taking on a more active role and the latter taking on a more passive role. This is unlike natural conversations which tend to be more collaborative. Another strength of this design is that it connected a word's salience to a speaker's communicative intention. This addition contributes to a more nuanced understanding of what different speakers remember from a conversation, which depends not only on what they uttered themselves but also on what they were interested in finding out. The present work shows that such factors should be incorporated more tightly in psycholinguistic research of dialogue to be able to develop a fuller understanding of the human capacity for dialogue.

Although the paradigm in its present form is still fairly artificial, it can easily be adjusted to encourage more naturalistic conversation. In Chapter 5, the answers that participants gave were selected for them beforehand and, as such, did not require the information-givers to be paying close attention to the questions; this was done for experimental control. In future work the information-giver could be allowed to choose the answer to the question from the two remaining objects on the screen. This would motivate the participants to pay attention to their partners' questions and perhaps engage them more fully in the experiments by asking them to make decisions. Similarly, the pre-selection of materials meant that participants had no intrinsic motivation or other reason to care about the answers to their questions. This could be manipulated experimentally by giving participants more flexibility in what they discuss (assuming that this will allow them to talk about topics they find interesting) or by making the answer relevant for the next turn in the task. Again, this would give the experimental more ecological validity.

Conclusion

What we remember from a conversation depends on our role in that conversation. In this thesis, I have demonstrated that we remember our own speech better than that of others because generating words and saying them aloud is beneficial to memory. That is, processes that occur during language production, but not during language comprehension, boost memory. Furthermore, I have shown that we remember the parts of a conversation that we consider important better than those we consider less important. This is the case not only when participating in a conversation but even when overhearing it. This thesis highlights the importance of studying both language production and language comprehension as active processes, especially during communication.

Bibliography

Agresti, A. (2007). *An introduction to categorical data analysis* (2nd ed.). Hoboken: New Jersey: John Wiley & Sons, Inc.

Almor, A., & Eimas, P. D. (2008). Focus and noun phrase anaphors in spoken language comprehension. *Language and Cognitive Processes*, *23*(2), 201-225. doi: 10.1080/01690960701330936

Audacity Team. (2014). *Audacity.* Audacity (R). Retrieved from `www.audacityteam.org`

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

Begg, I., Snider, A., Foley, F., & Goddard, R. (1989). The generation effect is no artifact: Generating makes works distinctive. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(5), 977-989. doi: http://dx.doi.org/10.1037/0278-7393.15.5.977

Benatar, A., & Clifton, C. (2014). Newness, givenness and discourse updating: Evidence from eye movements. *Journal of Memory and Language*, *71*(1), 1-16. doi: 10.1016/j.jml.2013.10.003

Bertsch, S., Pesta, B. J., Wiscott, R., & McDaniel, M. A. (2007). The generation effect: A meta-analytic review. *Memory & Cognition*, *35*(2), 201-210. doi: 10.3758/BF03193441

Birch, S. L., Albrecht, J. E., & Myers, J. L. (2000). Syntactic Focusing Structures Influence Discourse Processing. *Discourse Processes*, *30*(3), 285-304. doi: 10.1207/S15326950dp3003_4

Birch, S. L., & Garnsey, S. M. (1995). The Effect of Focus on Memory for Words in Sentences. *Journal of Memory and Language*, *34*(2), 232-267. doi: 10.1006/jmla.1995.1011

Birch, S. L., & Rayner, K. (1997). Linguistic focus affects eye movements during reading. *Memory & Cognition*, *25*(5), 653-660. doi: 10.3758/BF03211306

Birch, S. L., & Rayner, K. (2010). Effects of syntactic prominence on eye movements during reading. *Memory & Cognition*, *38*(6), 740-752. doi: 10.3758/MC.38.6.740

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.

Bodner, G. E., Jamieson, R. K., Cormack, D. T., McDonald, D.-L., & Bernstein, D. M. (2016). The production effect in recognition memory: Weakening strength can strengthen distinctiveness. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *70*(2), 93-98. doi: 10.1037/cep0000082

Boersma, P., & Weenink, D. (2016). *Praat: Doing phonetics by computer.*

Boersma, P., & Weenink, D. (2018). *Praat: Doing phonetics by computer.*

Borges, M. A., Stepnowsky, M. A., & Holt, L. H. (1977). Recall and recognition of words and pictures by adults and children. *Bulletin of the Psychonomic Society*, *9*(2), 113-114. doi: 10.3758/BF03336946

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, *105*(38), 14325-14329. doi: 10.1073/pnas.0803390105

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2013). Real-world objects are not represented as bound units: Independent forgetting of different object details from visual memory. *Journal of Experimental Psychology: General*, *142*(3), 791-808. doi: 10.1037/a0029649

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 717-726. doi: 10.1016/S0022-5371(72)80006-9

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1482. doi: 10.1037/0278-7393.22.6.1482

Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The Bank of Standardized Stimuli (BOSS), a New Set of 480 Normative Photos of Objects to Be Used as Visual Stimuli in Cognitive Research. *PLOS ONE*, *5*(5), e10773. doi: 10.1371/journal.pone.0010773

Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS) Phase II: 930 New Normative Photos. *PLOS ONE*, *9*(9), e106953. doi: 10.1371/journal.pone.0106953

Brown, A. S., Jones, E. M., & Davis, T. L. (1995). Age differences in conversational source monitoring. *Psychology and Aging*, *10*(1), 111. doi: 10.1037/0882-7974.10.1.111

Chomsky, N. (1971). Deep structure, surface structure, and semantic interpretation. In D. D. Steinberg & L. A. Jakobovits (Eds.), *An interdisciplinary reader in Philosophy, Linguistics and Psychology* (p. 183-216). New York: Cambridge University Press.

Conway, M. A., & Gathercole, S. E. (1987). Modality and long-term memory. *Journal of Memory and Language*, *26*(3), 341-361. doi: 10.1016/0749-596X(87)90118-5

Curran, T., & Doyle, J. (2011). Picture superiority doubly dissociates the ERP correlates of recollection and familiarity. *Journal of Cognitive Neuroscience*, *23*(5), 1247-1262. doi: https://doi.org/10.1162/jocn.2010.21464

Cutler, A., & Fodor, J. A. (1979). Semantic focus and sentence comprehension. *Cognition*, *7*(1), 49-59. doi: 10.1016/0010-0277(79)90010-6

DeCarlo, L. T. (1998). Signal detection theory and generalized linear models. *Psychological Methods*, *3*(2), 186-205. doi: http://dx.doi.org/10.1037/1082-989X.3.2.186

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition*, *42*(1), 287-314. doi: 10.1016/0010-0277(92)90046-K

Dew, I. T. Z., & Mulligan, N. W. (2008). The effects of generation on auditory implicit memory. *Memory & Cognition*, *36*(6), 1157–1167. doi: 10.3758/MC.36.6.1157

Dodson, C. S., & Schacter, D. L. (2001). "If I had said it I would have remembered it: Reducing false memories with a distinctiveness heuristic. *Psychonomic Bulletin & Review*, *8*(1), 155-161. doi: 10.3758/BF03196152

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191. doi: 10.3758/BF03193146

Fawcett, J. M., & Ozubko, J. D. (2016). Familiarity, but not recollection, supports the between-subject production effect in recognition memory. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *70*(2), 99-115. doi: 10.1037/cep0000089

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. doi: 10.1080/09658211.2012.693510

Fischer, N. M., Schult, J. C., & Steffens, M. C. (2015). Source and destination memory in face-to-face interaction: A multinomial modeling approach. *Journal of Experimental Psychology: Applied*, *21*(2), 195-204. doi: http://dx.doi.org/10.1037/xap0000046

Foraker, S., & McElree, B. (2007). The role of prominence in pronoun resolution: Active versus passive representations. *Journal of Memory and Language*, *56*(3), 357-383. doi: 10.1016/j.jml.2006.07.004

Forrin, N. D., Groot, B., & MacLeod, C. M. (2016). The d-Prime directive: Assessing costs and benefits in recognition by dissociating mixed-list false alarm rates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(7), 1090-1111. doi: 10.1037/xlm0000214

Forrin, N. D., Jonker, T. R., & MacLeod, C. M. (2014). Production improves memory equivalently following elaborative vs non-elaborative processing. *Memory*, *22*(5), 470-480. doi: 10.1080/09658211.2013.798417

Forrin, N. D., & MacLeod, C. M. (2018). This time it's personal: The memory benefit of hearing oneself. *Memory*, *26*(4), 574-579. doi: 10.1080/09658211.2017.1383434

Forrin, N. D., MacLeod, C. M., & Ozubko, J. D. (2012). Widening the boundaries of the production effect. *Memory & Cognition*, *40*(7), 1046-1055. doi: 10.3758/s13421-012-0210-8

Fraundorf, S. H., Watson, D. G., & Benjamin, A. S. (2010). Recognition memory reveals just how CONTRASTIVE contrastive accenting really is. *Journal of Memory and Language*, *63*(3), 367-386. doi: 10.1016/j.jml.2010.06.004

Gann, T. M., & Barr, D. J. (2014). Speaking from experience: Audience design as expert performance. *Language, Cognition and Neuroscience*, *29*(6), 744-760. doi: 10.1080/01690965.2011.641388

Ganushchak, L. Y., Konopka, A. E., & Chen, Y. (2014). What the eyes say about planning of focused referents during sentence formulation: A cross-linguistic investigation. *Frontiers in Psychology*, *5*. doi: 10.3389/fpsyg.2014.01124

Gardiner, J. M., & Hampton, J. A. (1985). Semantic memory and the generation effect: Some tests of the lexical activation hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(4), 732-741. doi: 10.1037/0278-7393.11.1-4.732

Gardiner, J. M., & Hampton, J. A. (1988). Item-specific processing and the generation effect: Support for a distinctiveness account. *The American Journal of Psychology*, *101*(4), 495-504. doi: 10.2307/1423227

Gaskell, M. G., Cairney, S. A., & Rodd, J. M. (2019). Contextual priming of word meanings is stabilized over sleep. *Cognition*, *182*, 109-126. doi: 10.1016/j.cognition.2018.09.007

Graesser, A. C., McMahen, C. L., & Johnson, B. K. (1994). Question asking and answering. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (p. 517-538). San Diego: Academic Press.

Halliday, M. A. K. (1967). Notes on transitivity and theme in English: Part 2. *Journal of Linguistics*, *3*(2), 199-244. doi: 10.1017/S0022226700016613

Hjelmquist, E. (1984). Memory for conversations. *Discourse Processes*, *7*(3), 321-336. doi: 10.1080/01638538409544595

Hoedemaker, R. S., Ernst, J., Meyer, A. S., & Belke, E. (2017). Language production in a shared task: Cumulative Semantic Interference from self- and other-produced context words. *Acta Psychologica*, *172*, 55-63. doi: 10.1016/j.actpsy.2016.11.007

Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language*, *32*(4), 421–445. doi: https://doi.org/10.1006/jmla.1993.1023

Hunt, R. R., & Worthen, J. B. (Eds.). (2006). *Distinctiveness and memory.*

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

Jacobs, C. L., Dell, G. S., Benjamin, A. S., & Bannard, C. (2016). Part and whole linguistic experience affect recognition memory for multi-word sequences. *Journal of Memory and Language*, *87*, 38-58. doi: 10.1016/j.jml.2015.11.001

Jamieson, R. K., Mewhort, D. J. K., & Hockley, W. E. (2016). A computational account of the production effect: Still playing twenty questions with nature. *Canadian Journal of Experimental Psychology/ Revue canadienne de psychologie experimentale*, *70*(2), 154–164. doi: 10.1037/cep0000081

Jarvella, R. J., & Collas, J. G. (1974). Memory for the intentions of sentences. *Memory & Cognition*, *2*(1), 185-188. doi: 10.3758/BF03197513

Järvikivi, J., Pyykkönen-Klauck, P., Schimke, S., Colonna, S., & Hemforth, B. (2014). Information structure cues for 4-year-olds and adults: Tracking eye movements to visually presented anaphoric referents. *Language, Cognition and Neuroscience*, *29*(7), 877-892. doi: 10.1080/01690965.2013.804941

Johns, C. L., Gordon, P. C., Long, D. L., & Swaab, T. Y. (2014). Memory availability and referential access. *Language, Cognition and Neuroscience*, *29*(1), 60-87. doi: 10.1080/01690965.2012.733014

Keenan, J. M., MacWhinney, B., & Mayhew, D. (1977). Pragmatics in memory: A study of natural conversation. *Journal of Verbal Learning and Verbal Behavior*, *16*(5), 549-560. doi: 10.1016/S0022-5371(77)80018-2

Kember, H., Choi, J., Yu, J., & Cutler, A. (2019). The Processing of Linguistic Prominence. *Language and Speech*, 0023830919880217. doi: 10.1177/0023830919880217

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. doi: 10.3758/BRM.42.3.643

Kinjo, H., & Snodgrass, J. G. (2000). Does the generation effect occur for pictures? *The American Journal of Psychology*, *113*(1), 95-121.

Kinoshita, S. (1989). Generation enhances semantic processing? The role of distinctiveness in the generation effect. *Memory & Cognition*, *17*(5), 563-571. doi: 10.3758/BF03197079

Knutsen, D., & Le Bigot, L. (2014). Capturing egocentric biases in reference reuse during collaborative dialogue. *Psychonomic Bulletin & Review*, *21*(6), 1590-1599. doi: 10.3758/s13423-014-0620-7

Kozminsky, E. (1977). Altering comprehension: The effect of biasing titles on text comprehension. *Memory & Cognition*, *5*(4), 482-490. doi: 10.3758/BF03197390

Levelt, W. J. M. (1989). The generation of messages. In *Speaking: From intention to articulation* (p. 107-160). Cambridge, MA: MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1-38.

Lowder, M. W., & Gordon, P. C. (2015). Focus takes time: Structural effects on reading. *Psychonomic Bulletin & Review*, *22*(6), 1733-1738. doi: 10.3758/s13423-015-0843-2

Machado, P., Romero, J., Nadal, M., Santos, A., Correia, J., & Carballal, A. (2015). Computerized measures of visual complexity. *Acta Psychologica*, *160*, 43-57. doi: 10.1016/j.actpsy.2015.06.005

MacLeod, C. M. (2011). I said, you said: The production effect gets personal. *Psychonomic Bulletin & Review*, *18*(6), 1197-1202. doi: 10.3758/s13423-011-0168-8

MacLeod, C. M., & Bodner, G. E. (2017). The production effect in memory. *Current Directions in Psychological Science*, *26*(4), 390-395. doi: 10.1177/0963721417691356

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), 671-685.

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

Mama, Y., & Icht, M. (2016). Auditioning the distinctiveness account: Expanding the production effect to the auditory modality reveals the superiority of writing over vocalising. *Memory*, *24*(1), 98-113. doi: 10.1080/09658211.2014.986135

McDaniel, M. A., & Bugg, J. M. (2008). Instability in memory phenomena: A common puzzle and a unifying explanation. *Psychonomic Bulletin & Review*, *15*(2), 237–255. doi: 10.3758/PBR.15.2.237

McKinley, G. L., Brown-Schmidt, S., & Benjamin, A. S. (2017). Memory for conversation and the development of common ground. *Memory & Cognition*, *45*(8), 1281-1294. doi: 10.3758/s13421-017-0730-3

Mckoon, G., Ratcliff, R., Ward, G., & Sproat, R. (1993). Syntactic Prominence Effects On Discourse Processes. *Journal of Memory and Language*, *32*(5), 593-607. doi: 10.1006/jmla.1993.1030

Miller, J. B. (1996). Memory for Conversation. *Memory*, *4*(6), 615-632. doi: 10.1080/741940999

Moreno-Martínez, F. J., & Montoro, P. R. (2012). An Ecological Alternative to Snodgrass & Vanderwart: 360 High Quality Colour Images with Norms for Seven Psycholinguistic Variables. *PLOS ONE*, *7*(5), e37527. doi: 10.1371/journal.pone.0037527

Morris, R. K., & Folk, J. R. (1998). Focus as a contextual priming mechanism in reading. *Memory & Cognition*, *26*(6), 1313-1322. doi: 10.3758/BF03201203

Nairne, J. S., Pusen, C., & Widner, R. L. (1985). Representation in the mental lexicon: Implications for theories of the generation effect. *Memory & Cognition*, *13*(2), 183-191. doi: 10.3758/BF03197011

Nairne, J. S., & Widner, R. L. (1987). Generation effects with nonwords: The role of test appropriateness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(1), 164-171. doi: 10.1037/0278-7393.13.1.164

Nelson, D. L., Reed, V. S., & McEvoy, C. L. (1977). Learning to order pictures and words: A model of sensory and semantic encoding. *Journal of Experimental Psychology: Human Learning and Memory*, *3*(5), 485-497. doi: 10.1037/0278-7393.3.5.485

Osaka, M., Nishizaki, Y., Komori, M., & Osaka, N. (2002). Effect of focus on verbal working memory: Critical role of the focus word in reading. *Memory & Cognition*, *30*(4), 562-571. doi: 10.3758/BF03194957

Ozubko, J. D., Hourihan, K. L., & MacLeod, C. M. (2012). Production benefits learning: The production effect endures and improves memory for text. *Memory*, *20*(7), 717-727. doi: 10.1080/09658211.2012.699070

Ozubko, J. D., & MacLeod, C. M. (2010). The production effect in memory: Evidence that distinctiveness underlies the benefit. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(6), 1543-1547. doi: 10.1037/a0020604

Ozubko, J. D., Major, J., & MacLeod, C. M. (2014). Remembered study mode: Support for the distinctiveness account of the production effect. *Memory*, *22*(5), 509–524. doi: 10.1080/09658211.2013.800554

Paivio, A., Rogers, T. B., & Smythe, P. C. (1968). Why are pictures easier to recall than words? *Psychonomic Science*, *11*(4), 137-138. doi: 10.3758/BF03331011

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

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

Powell, M. J. D. (2009). *The BOBYQA algorithm for bound constrained optimization without derivatives* (Tech. Rep. No. Cambridge NA Report NA2009/06). Cambridge: University of Cambridge.

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language*, *60*(4), 437–447. doi: 10.1016/j.jml.2009.01.004

Quinlan, C. K., & Taylor, T. L. (2013). Enhancing the production effect in memory. *Memory*, *21*(8), 904-915. doi: 10.1080/09658211.2013.766754

R Core Team. (2017). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing. Vienna, Austria.

R Core Team. (2018). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing. Vienna, Austria.

R Core Team. (2019). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing. Vienna, Austria.

Raven, J. (1998). *Manual for Raven's advanced progressive matrices*. Oxford, UK: Oxford Psychologists Press.

Richler, J. J., Palmeri, T. J., & Gauthier, I. (2013). How does using object names influence visual recognition memory? *Journal of Memory and Language*, *68*, 10–25. doi: http://dx.doi.org/10.1016/j.jml.2012.09.001

Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, *57*(3), 348–379. doi: 10.1016/j.jml.2007.03.002

Rosson, M. B. (1983). From SOFA to LOUCH: Lexical contributions to pseudoword pronunciation. *Memory & Cognition*, *11*(2), 152-160. doi: 10.3758/BF03213470

Sanford, A. J. S., Price, J., & Sanford, A. J. (2009). Enhancement and suppression effects resulting from information structuring in sentences. *Memory & Cognition*, *37*(6), 880–888. doi: 10.3758/MC.37.6.880

Sauermann, A., Filik, R., & Paterson, K. B. (2013). Processing contextual and lexical cues to focus: Evidence from eye movements in reading. *Language and Cognitive Processes*, *28*(6), 875-903. doi: 10.1080/01690965.2012.668197

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

Schweitzer, A., & Lewandowski, N. (2013). Convergence of articulation rate in spontaneous speech. In *Proceedings of the 14th annual conference of the international speech communication association (interspeech 2013, lyon)* (pp. 525–529).

Schweitzer, A., & Lewandowski, N. (2014). Social factors in convergence of f1 and f2 in spontaneous speech. In *Proceedings of the 10th interna-*

tional seminar on speech production, cologne.

Simonsohn, U. (2013). *Titleology: Some facts about titles.* Retrieved 2020-04-04, from `http://datacolada.org/9`

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, *4*(6), 592-604. doi: http://dx.doi.org/10.1037/0278-7393.4.6.592

Slamecka, N. J., & Katsaiti, L. T. (1987). The generation effect as an artifact of selective displaced rehearsal. *Journal of Memory and Language*, *26*(6), 589-607. doi: 10.1016/0749-596X(87)90104-5

Smith, M. C., & Magee, L. E. (1980). Tracing the time course of picture–word processing. *Journal of Experimental Psychology: General*, *109*(4), 373-392. doi: 10.1037/0096-3445.109.4.373

Stafford, L., & Daly, J. A. (1984). Conversational Memory. *Human Communication Research*, *10*(3), 379-402. doi: 10.1111/j.1468-2958.1984.tb00024.x

Sturt, P., Sanford, A. J., Stewart, A., & Dawydiak, E. (2004). Linguistic focus and good-enough representations: An application of the change-detection paradigm. *Psychonomic Bulletin & Review*, *11*(5), 882-888. doi: 10.3758/BF03196716

Theios, J., & Muise, J. G. (1977). The word identification process in reading. *Cognitive theory*, *2*, 289-321.

van Casteren, M., & Davis, M. H. (2007). Match: A program to assist in matching the conditions of factorial experiments. *Behavior Research Methods*, *39*(4), 973-978. doi: 10.3758/BF03192992

Van Der Wege, M. M. (2009). Lexical entrainment and lexical differentiation in reference phrase choice. *Journal of Memory and Language*, *60*(4), 448-463. doi: 10.1016/j.jml.2008.12.003

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

Ward, P., & Sturt, P. (2007). Linguistic focus and memory: An eye movement study. *Memory & Cognition*, *35*(1), 73-86. doi: 10.3758/BF03195944

Weldon, M. S., & Roediger, H. L. (1987). Altering retrieval demands reverses the picture superiority effect. *Memory & Cognition*, *15*(4), 269-280. doi: 10.3758/BF03197030

Wright, D. B., Horry, R., & Skagerberg, E. M. (2009). Functions for traditional and multilevel approaches to signal detection theory. *Behavior Research Methods*, *41*(2), 257-267. doi: 10.3758/BRM.41.2.257

Yang, X., Zhang, X., Wang, C., Chang, R., & Li, W. (2017). The Interplay between Topic Shift and Focus in the Dynamic Construction of Discourse Representations. *Frontiers in Psychology*, *8*. doi: 10.3389/fpsyg.2017.02184

Yoon, S. O., Benjamin, A. S., & Brown-Schmidt, S. (2016). The historical context in conversation: Lexical differentiation and memory for the discourse history. *Cognition*, *154*, 102-117. doi: https://doi.org/10.1016/j.cognition.2016.05.011

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. doi: 10.1080/09658211.2018.1510966

Zormpa, E., Meyer, A. S., & Brehm, L. E. (2019). Slow naming of pictures facilitates memory for their names. *Psychonomic Bulletin & Review*. doi: 10.3758/s13423-019-01620-x

Nederlandse samenvatting

Het merendeel van wat we leren, leren we via taal. Soms leren we iets als we in ons eentje een boek lezen of als we naar het nieuws luisteren, soms leren we als we in gesprek zijn met anderen. Je zou denken dat we dus bijzonder goed zijn in het onthouden van taal die we gehoord of gelezen hebben. Toch lijkt dat niet het geval te zijn, tenminste niet in experimenten met één persoon. Decennia aan geheugenonderzoek hebben aangetoond dat we beter zijn in het onthouden van woorden die we zelf produceren, dan van woorden die we horen of lezen. Om precies te zijn onthouden we woorden die we *bedenken* beter dan woorden die we *lezen* en onthouden we woorden die we *hardop* zeggen beter dan woorden die we *in ons hoofd* zeggen. Mensen zover krijgen dat ze een bepaald woord zeggen is verrassend moeilijk (mensen kunnen heel onvoorspelbaar zijn!) en dus heeft veel van dit onderzoek eruit bestaan deelnemers een woord te tonen en ze te vragen om het tegenovergestelde woord te bedenken, terwijl ze ook de eerste letter te zien krijgen van het woord dat ze moeten bedenken. Dit is een effectieve taak, maar het staat ver af van hoe we woorden bedenken in de echte wereld, behalve misschien bij het oplossen van kruiswoordpuzzels. Zouden de uitkomsten van deze onderzoeken veranderen als we taal testen op een manier die meer lijkt op hoe we spreken in het dagelijks leven?

In deze dissertatie onderzocht ik het geheugen voor spreken en luisteren met taal die meer natuurlijk is. Ik deed dit zowel in situaties waar één persoon een enkel woord spreekt, als in situaties waarin twee personen een langer gesprek hebben. De taken die ik gebruikt heb, waren ook kunstmatig, maar ze bootsen normaal taalgebruik beter na. In éénpersoonsexperimenten vroeg ik deelnemers om plaatjes te benoemen. In tweepersoonsexperimenten vroeg ik deelnemers om de namen van plaatjes die ze zagen te gebruiken in vraag-en-antwoord interacties. Mijn doel was om te onderzoeken hoe het geheugen beïnvloed wordt door de handeling van het spreken zelf en door onderliggende processen, maar daarnaast ook hoe het geheugen beïnvloed wordt door het belang van wat een ander zegt.

In **Hoofdstuk 2** beschrijf ik een onderzoek waarin ik test of de twee bevindingen die ik eerder noemde (dat zowel het bedenken als het hardop zeggen van woorden het geheugen bevorderen) ook gevonden kunnen worden in een taak waarin plaatjes benoemd moeten worden (een plaatjesbenoemtaak). Om dat te bereiken maakte ik een versie van de plaatjesbenoemtaak met twee aanpassingen. Ten eerste waren de plaatjes altijd gelabeld. Dit label kon de naam van het plaatje zijn, wat betekent dat deelnemers niet over de naam na hoefden te denken, maar het kon ook een combinatie van onleesbare, door elkaar gegooide letters zijn, waardoor de deelnemers zelf de naam van het plaatje moesten bedenken. Ten tweede waren de plaatjes altijd omgeven door een vierkant: een groen vierkant als deelnemers het plaatje hardop moesten benoemen en een rood vierkant als ze het alleen in hun hoofd moesten benoemen. Ze benoemden 128 plaatjes op deze manier, waarna ze 20 minuten later de plaatjes opnieuw zagen, gemengd met nieuwe plaatjes, en moesten aangeven of ze de plaatjes eerder gezien hadden. Ik vond dat deelnemers beter waren in het herkennen van plaatjes die ze eerder benoemd hadden, dan in plaatjes die ze niet benoemd hadden. Ze konden zich ook beter de plaatjes herinneren die ze hardop benoemd hadden dan de plaatjes die ze alleen in hun hoofd benoemd hadden. Dit toont dat de geheugeneffecten waar ik in geïnteresseerd was, gevonden kunnen worden met een meer natuurlijke manier om deelnemers woorden te laten produceren.

In **Hoofdstuk 3** keek ik in meer detail naar de bevinding dat mensen woorden beter onthouden als ze ze zelf bedenken, dan als ze ze lezen of horen, en onderzocht op welk 'niveau' deze geheugenverbetering plaatsvindt. Onthouden mensen het plaatje dat ze benoemd hebben? Of onthouden ze het woord dat ze gezegd hebben? Om deze vraag te beantwoorden, herhaalde ik het vorige experiment, maar met een paar aanpassingen. Deelnemers benoemden plaatjes altijd hardop en in plaats van een label te *lezen*, *hoorden* ze ofwel de naam van het plaatje, ofwel een ongerelateerd woord achterstevoren afgespeeld, ofwel een ongerelateerd woord normaal afgespeeld. Daarna vroeg ik de deelnemers om niet de plaatjes die ze hadden benoemd te herkennen, maar de namen van de plaatjes die ze gezien hadden. Ik vond dat mensen nog steeds beter waren in het herkennen van woorden die ze zelf benoemd hadden, dan in woorden die ze simpelweg herhaald hadden, wat aantoont dat niet alleen het geheugen voor plaatjes, maar ook het geheugen voor woorden verbetert. Ik vond ook dat wanneer mensen er lang over deden om een plaatje te benoemen, dit hun geheugen meer verbeterde dan als ze een plaatje snel benoemden. De ontdekking dat dit alleen gebeurde met woorden die deelnemers zelf bedacht (in plaats van herhaald) hadden, toont aan dat de handeling van het benoemen van een plaatje het geheugen beïnvloedt.

In het onderzoek beschreven in **Hoofdstuk 4** verlegde ik mijn aandacht van geïsoleerde woorden naar langere spraakfragmenten, vragen en antwoorden om precies te zijn. Ik wilde uitzoeken of delen van een interactie die belangrijk zijn ook beter worden onthouden. In deze situatie is het krijgen van een antwoord de reden dat iemand een vraag stelt, wat het antwoord belangrijk maakt. Dit experiment was anders dan de vorige experimenten, niet alleen omdat ik korte dialogen gebruikte in plaats van geïsoleerde woorden, maar ook omdat deelnemers niet meer zelf spraken. In plaats daarvan zagen ze drie plaatjes op het scherm terwijl ze luisterden naar een vraag-en-antwoord interactie waarin twee van de plaatjes werden genoemd. Toen ik deelnemers vroeg om de namen van alle plaatjes die ze gezien hadden te herkennen waren ze beter in het herkennen van woorden die ze gehoord hadden in de antwoorden, dan woorden die ze gehoord hadden in de vragen.

Ik combineer beide stromen van onderzoek in mijn dissertatie in **Hoofdstuk 5**, waarin ik vraag hoe spraakproductie en het belang van wat er gezegd wordt het geheugen beïnvloeden. In dit experiment zagen twee deelnemers drie plaatjes op een scherm en stelden en beantwoordden ze vragen over hoe de plaatjes verschoven moesten worden. De ene deelnemer vroeg bijvoorbeeld "Wat hoort er naast het schilderij?" waarop de ander antwoordde "De geit!". Door de COVID-19-uitbraak kon de dataverzameling voor dit experiment niet voltooid worden, maar de tussentijdse resultaten komen overeen met mijn eerdere bevindingen. Dat wil zeggen, mensen onthielden wat zij zelf zeiden het best, maar als ze gevraagd werd om een woord te herkennen dat ze gehoord hadden waren ze daar beter in als dat woord in een antwoord was voorgekomen dan als het woord in een vraag was voorgekomen.

Kortom, ik heb aangetoond dat mensen beter onthouden wat ze zeggen, dan wat ze horen. Ik heb ook laten zien dat het proces van het bedenken van een woord daar belangrijk voor lijkt te zijn–hoe langer het proces duurt, hoe beter het woord onthouden wordt. Mensen onthouden wat ze zeggen niet alleen goed wanneer ze tegen zichzelf praten, maar ook wanneer ze tegen een ander praten. Dat betekent echter niet dat ze alles wat ze horen vergeten–als iets belangrijk is dan wordt het goed onthouden; het antwoord op een vraag wordt bijvoorbeeld goed onthouden, in verhouding tot de vraag. Dit werk toont het belang van het bestuderen van meer natuurlijk spraakgebruik en het opent deuren voor gezamenlijk onderzoek naar taal en geheugen.

Translated from English by Jeroen van Paridon

English summary

We do most of our learning through language, sometimes by ourselves when reading a book or listening to the news, and sometimes with others through conversation. One would think, then, that we are especially good at remembering the language that we comprehend (what we hear or read). However, that is not the case, at least not in studies with one person. Decades of memory research have shown that we are better at remembering words we produce ourselves than words we comprehend. More specifically, we remember the words we *come up with* better than the words we *read* and we remember the words we say *aloud* better than the words we say *in our heads*. Getting people to produce a certain word is surprisingly hard (people can be very unpredictable!), so much of this research has relied on showing participants a word and asking them to come up with its opposite while also showing them the first letter for the word they need to come up with. This is an effective task, but it is a far cry from how we produce words in real life, except when solving crosswords. So might these findings change if the language we test is a little closer to how we speak in everyday life?

In this thesis, I explored memory for speaking and listening with language that is a little more natural. I did this both in situations where one person says a single word and in situations where two people have a longer exchange. The tasks I used are also artificial, but they mimic normal language use a little better. In one-person experiments, I asked participants to name pictures. In two-person experiments, I asked participants to use the names of pictures they saw in question-answer exchanges. My goal was to see how memory is affected by the act of speaking itself and its component processes, but also how it is affected by the importance of what someone else is saying.

In **Chapter 2**, I describe a study in which I test if the two findings I mentioned previously (that both coming up with words and saying them out loud improve memory) can be found with a picture naming task. To do

that, I created a version of a picture naming experiment with two modifications. Firstly, pictures always had a label on them. This label could be the name of the picture, meaning participants did not have to think of the name themselves, or jumbled up letters that made no sense, in which case participants did have to think of a name themselves. Secondly, pictures always had a square around them: the square was green if participants had to name the picture aloud and red if they had to name it silently in their heads. They named 128 pictures this way and, after 20 minutes, they saw the pictures again, intermixed with some new pictures, and had to decide if they had seen these pictures before. I found that participants were better at recognising the pictures they had named than the ones they had not. They also remembered the pictures they had named aloud better than the pictures they had named silently. This shows that the memory effects in which I was interested can be detected with a more natural way of producing words.

In **Chapter 3**, I zoomed in on the finding that people remember the words they come up with better than the words they read or hear and asked at what 'level' this memory improvement happens. Do people remember the picture they named? Or do they remember the word they said? To answer this question, I ran a similar experiment to the previous one with a few changes. Participants always named the pictures aloud and, instead of *reading* words, they *heard* either the name of the picture, an unrelated word played backwards, or an unrelated word played normally. I then asked participants to recognise not the pictures they named, but the names of the pictures they saw. I found that people were still better at recognising words they named themselves than words they just repeated, showing that memory for words improves too, not just memory for pictures. I also found that when people took a long time to name a picture, this improved their memory relative to when they took little time to name a picture. The finding that this only happened with words participants had thought of themselves (not repeated) shows that the act of naming a picture influences memory.

In the study described in **Chapter 4**, I turned my attention from isolated words to longer stretches of speech, more specifically question-answer pairs. Here I wanted to find out if parts of an exchange that are more important are remembered better. In this case, the reason someone asks a question is to find out the answer, making the answer more important. This experiment was different to the previous ones not only because I used short dialogues instead of isolated words, but also because participants were no longer speaking themselves. Instead, they saw three pictures on the screen while listening to a question-answer exchange in which two of the pictures were mentioned. When I asked participants to recognise the names of all the pictures they had seen, they were more accurate with words that they had heard in the answers than words they had heard in the questions.

I combine the two streams of work in my thesis in **Chapter 5**, where I ask how speech production and the importance of what is being said influence memory. In this experiment, two participants saw three pictures on a screen and asked and answered questions about how the pictures should move. For example, one participant would ask "What should go next to the painting?" and the other would answer "The goat!". Due to the COVID-19 outbreak, data collection could not be completed but the results so far are in agreement with my previous findings. That is, people remembered what they said themselves best, but when they were asked to recognise a word they had heard, they were more accurate when that word had appeared in an answer than when it had appeared in a question.

In summary, I have shown that people remember what they say better than what they hear. I have also shown that the process during which people come up with words seems to be important for this benefit–the longer this process takes, the better the word is remembered. People remember what they say well not only when they are talking to themselves, but also when they are talking to somebody else. That does not mean, however, that they forget everything they hear–when something is considered important, for example an answer to a question, people remember the answer quite well relative to the question. This work shows the importance of studying more naturalistic speech and opens up new possibilities in the joint study of memory and language.

Ελληνική περίληψη

Η γλώσσα αποτελεί το βασικό μέσο μάθησης, είτε αυτή επιτυγχάνεται μέσω της ανάγνωσης βιβλίων είτε μέσω συζητήσεων και της παρακολούθησης ειδήσεων. Θα φανταζόταν, λοιπόν, κανείς ότι ο άνθρωπος έχει συνηθίσει να θυμάται όσα καταλαβαίνει είτε προφορικώς είτε γραπτώς. Στην πραγματικότητα όμως αυτό δεν ισχύει, τουλάχιστον όχι σε μελέτες που έχουν πραγματοποιηθεί με τη συμμετοχή ενός ατόμου. Εδώ και δεκαετίες, η επιστημονική μελέτη της μνήμης αποδεικνύει ότι θυμόμαστε καλύτερα τις λέξεις που παράγουμε οι ίδιοι παρά τις λέξεις τις οποίες κατανοούμε, αυτές που ακούμε δηλαδή ή διαβάζουμε. Πιο συγκεκριμένα, θυμόμαστε καλύτερα τις λέξεις τις οποίες σκεφτόμαστε οι ίδιοι σε σχέση με τις λέξεις τις οποίες διαβάζουμε. Επίσης, θυμόμαστε καλύτερα τις λέξεις τις οποίες προφέρουμε φωναχτά παρά τις λέξεις τις οποίες σκεφτόμαστε από μέσα μας. Μια από τις δυσκολίες που αντιμετωπίζονται σε τέτοιου είδους μελέτες είναι ο τρόπος με τον οποίο μπορεί κανείς να κάνει τους συμμετέχοντες να προφέρουν μια συγκεκριμένη λέξη (είναι απρόβλεπτοι οι άνθρωποι!). Γι' αυτόν τον λόγο, σε τέτοιου είδους μελέτες οι συμμετέχοντες συνήθως καλούνται να σκεφτούν το συνώνυμο ή το αντώνυμο μιας λέξης ενώ ταυτόχρονα έχουν στη διάθεσή τους το πρώτο γράμμα της ίδιας λέξης. Αυτό, φυσικά, ουδεμία σχέση έχει με τον καθημερινό λόγο, με εξαίρεση την περίπτωση της λύσης σταυρόλεξων. Θα μπορούσαν όμως αυτά τα ευρήματα να ήταν διαφορετικά αν η γλώσσα που χρησιμοποιείται στις έρευνες ήταν πιο κοντά στον καθημερινό λόγο;

Σκοπός της παρούσας διδακτορικής διατριβής είναι η διερεύνηση της μνήμης κατά την ομιλία και την ακρόαση, μέσω της χρήσης πιο φυσικής γλώσσας σε σχέση με αυτή που έχει χρησιμοποιηθεί σε προϋπάρχουσες έρευνες. Έτσι, μελετήθηκαν περιπτώσεις στις οποίες ένας συμμετέχων πρόφερε μία λέξη αλλά και περιπτώσεις στις οποίες δύο συμμετέχοντες συνομιλούσαν. Παρόλο που τα πειράματα που πραγματοποιήθηκαν ήταν αρκετά τεχνητά, η ομιλία των συμμετεχόντων βρίσκεται πιο κοντά στον καθημερινό λόγο. Στα πειράματα με ένα άτομο, οι συμμετέχοντες έβλεπαν φωτογραφίες αντικειμένων τα οποία κλήθηκαν να κατονομάσουν. Στα πειράματα με δύο άτομα, οι συμμετέχοντες χρησιμοποιούσαν τα ονόματα αντικειμένων που έβλεπαν σε φωτογραφίες υπό τη μορφή ερωτοαπαντήσεων. Στόχο αποτέλεσε η διερεύνηση του τρόπου με τον οποίο επηρεάζεται η μνήμη πρώτον από την ίδια την ομιλία και τις συνιστώσες της διαδικασίες και δεύτερον των όσων εκφράζει ο συνομιλητής.

Στο Κεφάλαιο 2, περιγράφεται η μελέτη στην οποία ερευνάται το κατά πόσον μια άσκηση κατονομασίας εικόνων μπορεί να οδηγήσει στα δύο ευρήματα που προαναφέρθηκαν –αν το να βρίσκει από μόνος του κάποιος μία λέξη και να την προφέρει δυνατά μπορούν να οδηγήσουν στη βελτίωση της μνήμης. Για τον λόγο αυτό κρίθηκαν απαραίτητες δύο αλλαγές στο κλασικό πείραμα στο οποίο οι συμμετέχοντες κατονομάζουν τα αντικείμενα που βλέπουν. Πρώτον, όλες οι φωτογραφίες είχαν πάνω τους μια κάρτα στην οποία ορισμένες φορές ήταν γραμμένο το όνομα του αντικειμένου της φωτογραφίας –οπότε και οι συμμετέχοντες δεν χρειαζόταν να σκεφτούν το όνομα του αντικειμένου– ενώ άλλες φορές ήταν γράμματα κομμένα σε μικρά κομματάκια και τοποθετημένα με μπερδεμένη σειρά –οπότε κι οι συμμετέχοντες έπρεπε να βρουν το κατάλληλο όνομα μόνοι τους. Δεύτερον, όλες οι φωτογραφίες περιβάλλονταν από ένα τετράγωνο: όταν το τετράγωνο ήταν πράσινο, οι συμμετέχοντες κατονόμαζαν το αντικείμενο δυνατά ενώ όταν ήταν κόκκινο το διάβαζαν από μέσα τους. Με αυτόν τον τρόπο, οι συμμετέχοντες κατονόμασαν συνολικά 128 φωτογραφίες και 20 λεπτά αργότερα είδαν πάλι τις ίδιες φωτογραφίες μαζί με άλλες 128 καινούριες φωτογραφίες ενώ τους ζητήθηκε να αποφασίσουν ποιες από τις φωτογραφίες είχαν δει προηγουμένως. Αποδείχθηκε ότι οι συμμετέχοντες θυμόνταν καλύτερα τις φωτογραφίες που είχαν κατονομάσει οι ίδιοι σε σχέση με εκείνες των οποίων το όνομα είχαν διαβάσει. Επίσης, αποδείχθηκε ότι θυμόνταν καλύτερα τις φωτογραφίες που είχαν κατονομάσει δυνατά σε σχέση με εκείνες που είχαν διαβάσει από μέσα τους. Αυτά τα αποτελέσματα επιβεβαιώνουν ότι τα μνημονικά φαινόμενα που ενδιαφέρουν την παρούσα διατριβή εντοπίζονται και στην περίπτωση χρήσης πιο καθημερινού λόγου.

Στο Κεφάλαιο 3, η προσοχή της μελέτης στράφηκε στο εύρημα σύμφωνα με το οποίο θυμόμαστε καλύτερα τις λέξεις τις οποίες σκεφτόμαστε οι ίδιοι σε σχέση με αυτές που διαβάζουμε ή ακούμε, με σκοπό να προσδιοριστεί το «επίπεδο» βελτίωσης της ανθρώπινης μνήμης. Δηλαδή, είναι η ίδια η φωτογραφία αυτό που θυμόμαστε καλύτερα; Ή μήπως είναι η λέξη που προφέρουμε; Προκειμένου, λοιπόν, να βρεθεί η απάντηση σε αυτό το ερώτημα, πραγματοποιήθηκε ένα πείραμα παρόμοιο με το προηγούμενο αλλά στο οποίο εφαρμόστηκαν δύο αλλαγές. Σ' αυτό το πείραμα, οι συμμετέχοντες κατονόμαζαν το αντικείμενο πάντα δυνατά και αντί να διαβάζουν κάποια λέξη στην κάρτα, άκουγαν ή το όνομα του αντικειμένου ή μία άσχετη λέξη παιγμένη ανάποδα ή μια άσχετη λέξη παιγμένη κανονικά. Έπειτα οι συμμετέχοντες κλήθηκαν να αναγνωρίσουν τα ονόματα των αντικειμένων που είχαν δει προηγουμένως και όχι τα ίδια τα αντικείμενα. Τα αποτελέσματα δείχνουν ότι οι συμμετέχοντες θυμόνταν τα ονόματα των αντικειμένων που είχαν ονομάσει οι ίδιοι καλύτερα από τα ονόματα των αντικειμένων που είχαν ακούσει και απλά επανέλαβαν. Αυτό σημαίνει ότι κατά την παραγωγή μιας λέξης από μια φωτογραφία, βελτιώνεται η μνήμη όχι μόνο για την φωτογραφία αλλά και για την ίδια την λέξη. Τα αποτελέσματα δείχνουν ακόμη ότι όσο περισσότερη ώρα χρειάζονταν οι συμμετέχοντες για να κατονομάσουν ένα αντικείμενο, τόσο καλύτερα θυμόνταν το όνομα του αντικειμένου. Το γεγονός ότι αυτό συνέβαινε μόνο όταν οι συμμετέχοντες έβρισκαν την λέξη από μόνοι τους –και όχι όταν την επαναλάμβαναν– αποδεικνύει ότι η ίδια η πράξη της κατονομασίας ενός αντικειμένου επηρεάζει την μνήμη.

Στη μελέτη που περιγράφεται στο Κεφάλαιο 4, η προσοχή της έρευνας στράφηκε από τη μελέτη απομονωμένων λέξεων στη μελέτη πιο εκτεταμένης ομιλίας και συγκεκριμένα στη μελέτη ερωτοαπαντήσεων. Σκοπός αυτής της μελέτης ήταν να διαπιστωθεί κατά πόσο θυμόμαστε καλύτερα κάτι που θεωρείται σημαντικό σε μια συναναστροφή. Στην προκειμένη περίπτωση, θεωρήθηκε ότι εφόσον, συνήθως, ο λόγος που θέτουμε ερωτήσεις είναι γιατί αποσκοπούμε στην απάντηση, αυτή θα είναι πιο σημαντική από την ερώτηση. Αυτό το πείραμα διαφέρει από τα προηγούμενα όχι μόνο επειδή χρησιμοποιήθηκαν σύντομοι διάλογοι αντί απλές λέξεις αλλά και επειδή σε αυτή την περίπτωση οι συμμετέχοντες δεν μιλούσαν. Αντ' αυτού, έβλεπαν τρεις φωτογραφίες στην οθόνη ενός υπολογιστή ενώ ταυτόχρονα άκουγαν έναν σύντομο διάλογο στον οποίο η ερώτηση ανέφερε μια από τις φωτογραφίες και η απάντηση μια δεύτερη. Αργότερα, οι συμμετέχοντες κλήθηκαν να αναγνωρίσουν τα όνομα των φωτογραφιών που είχαν δει. Όπως αποδείχθηκε, οι συμμετέχοντες θυμόνταν τις λέξεις που άκουσαν στις απαντήσεις καλύτερα από ότι τις λέξεις που άκουσαν στις ερωτήσεις.

Στο Κεφάλαιο 5 συνδυάστηκαν οι δύο κατευθύνσεις που περιγράφονται στην παρούσα διατριβή, με σκοπό να διαπιστωθεί σε ποιο βαθμό τα όσα λέγονται σε μία συζήτηση και η σημασία τους επηρεάζουν τη μνήμη. Σε αυτό το πείραμα, δύο συμμετέχοντες έβλεπαν τρεις φωτογραφίες στην οθόνη ενός υπολογιστή και έθεταν και απαντούσαν ερωτήσεις για το πώς θα έπρεπε να μετακινηθούν αυτές οι φωτογραφίες. Για παράδειγμα, ένας συμμετέχων μπορεί να ρωτούσε «Τι πρέπει να πάει δίπλα στον πίνακα ζωγραφικής;» και ο άλλος να απαντούσε «Η κατσίκα!». Λόγω της πανδημίας του COVID-19, η συλλογή όλων των απαραίτητων δεδομένων δεν κατέστη δυνατή, ωστόσο, τα αποτελέσματα μέχρι τώρα επιβεβαιώνουν τα προηγούμενα ευρήματα. Δηλαδή, οι συμμετέχοντες θυμόνταν καλύτερα τι είπαν αυτοί οι ίδιοι ενώ όταν έπρεπε να αναγνωρίσουν λέξεις τις οποίες είχαν ακούσει, θυμόνταν πολύ καλύτερα τις λέξεις που είχαν αναφερθεί σε απάντηση σε σχέση με αυτές που είχαν αναφερθεί σε ερώτηση.

Κλείνοντας, στην παρούσα διατριβή αποδείχθηκε ότι ο άνθρωπος θυμάται αυτά που λέει καλύτερα σε σχέση με αυτά που ακούει. Αποδείχθηκε, επίσης, ότι η διαδικασία μέσω της οποίας βρίσκουμε τις λέξεις που θέλουμε να χρησιμοποιήσουμε παίζει σημαντικό ρόλο σ΄ αυτό το μνημονικό όφελος· όσο περισσότερο παίρνει αυτή η διαδικασία, τόσο καλύτερα θυμόμαστε τη λέξη που είπαμε. Επίσης, θυμόμαστε καλά αυτά που λέμε όχι μόνο όταν μιλάμε μόνοι μας αλλά και όταν μιλάμε με άλλους. Ωστόσο αυτό δεν σημαίνει ότι ξεχνάμε ό,τι ακούμε. Αντιθέτως, όταν ακούμε κάτι που θεωρούμε σημαντικό όπως στην περίπτωση της απάντησης σε μια ερώτηση, τότε θυμόμαστε την απάντηση αρκετά καλύτερα σε σχέση με την ερώτηση. Η παρούσα έρευνα αποδεικνύει τη σημασία της μελέτης μιας περισσότερο φυσικής γλώσσας και ανοίγει καινούριες προοπτικές στην κοινή έρευνα της γλώσσας και της μνήμης.

Acknowledgements

It was the best of times, it was the worst of times. Or something like that. Doing a PhD is hard (who would have thought), but I was very lucky to have an amazing group of people around me to share the fun bits of doing a PhD and to help me through the not-so-fun bits of doing a PhD.

None of this would have been possible without a wonderful supervisory team. Thank you Antje for giving me the chance to do a PhD when I had almost given up on the idea. Thank you also for helping me establish my project, and then another one after the first one didn't work out. I've greatly enjoyed coming up with experiments with you, particularly the part that involved drawing things that look like cats. Thank you especially for your guidance in the last few months through the craziness of COVID. I really appreciate it.

Renske, you made me feel welcome at the MPI from our first meeting. Thank you for teaching me about experimental design and data analysis. I always felt comfortable to come knock on your door (sorry Suzanne!), bewildered and confused by SPSS syntax or some unruly eye-tracker.

Laurel, I could not have been luckier to have you as my "new" supervisor. Your intelligence and kindness were exactly what I needed to get me through this PhD. Thank you for patiently teaching me how to write academic articles and code. It may have taken approximately four hours, but you taught me how to write my first function—and one for power analyses to boot! Thank you for always being there to answer silly R questions, for calming me down every time I freaked out about my analyses and, crucially, for being a reliable source of snacks and cat pictures.

Thank you also to the members of the manuscript committee Herbert Schriefers, Eva Belke, and Andreas Mädebach for reading and evaluating my thesis.

I am so grateful for my παρανυφάκια. Greta, I can't quite express how happy I am we started our PhDs together. Not only because it meant we were going through phases of excitement and disillusionment more or less together, but also because I can't really imagine being at the MPI without you. Thank you for all our long chats (a non-trivial amount of which started with me coming to your office for "a 5 minute break") and for always making me feel understood. I could not have asked for a better friend.

Joe and Jeroen, you both intimidated the hell out of me when I first met you, but you also quickly made me feel like I belonged at the MPI. Thank you so much for that, I am truly lucky I can call you my friends. Joe, you are someone I always felt I could turn to for help, in and out of work. I had a great time sharing an office with you and, briefly, a flat. Thank you for your help with R and Illustrator, for parties and amazing meals, and for carrying about half of my belongings over the Channel. Thank you also for all your help and excellent suggestions for my thesis cover.

Jeroen, I am so happy that you were my office-mate for almost all of my PhD. Thank you for the endless supply of coffee and for making 374 a fun, if slightly strange, office to be in. Thank you also for translating my summary and for the (many) times you helped me out with your Python wizardry. Most importantly, thank you for giving me access to your massive TV and The Professor.

Federica, thank you for always being a considerate office mate and I'm sorry for not returning the favour. You can now enjoy a dark office, little vampire!

Thank you to all the members of the Psychology of Language department over the years: Aitor, Alastair, Andrea, Evelyn, Falk, Fan, Florian, Hans Rutger, Jieying, Johanne, Markus, Merel M., Merel W., Shiri, Suzanne, Thy, Will, and Zeshu for fun lunches and engaging lab and cluster meetings.

Thank you especially Amie for being one of the loveliest people I know and for introducing me to the world of open science. Caitlin and Sophie, our virtual coffee chats in the last few months really helped me keep my sanity, such as it is. Thank you Phillip for your contribution to my transformation into an absolute coffee snob but mostly for kicking me out of the office when I really needed someone to do that. Sara, thank you for being such a wonderful and dynamic person in and out of work. I used to be apprehensive about organising stuff, but that was not really an option around you. Miguel, thank you for always having seen the movies I wanted to chat about. Thank you Saoradh for being a great person to start a PhD with and for a lovely ESCoP. Limor, thank you for forcing vegan pancakes

on me every now and then. Nina, thank you for all the goat and unicorn pictures over the years.

A huge thank you to Annelies and all the amazing PoL student assistants for all the translations, judgements and everything else I've asked for over the years. Thank you especially to Dennis for lending me his voice and withstanding a seemingly endless stream of "What should go next to the ..." questions. Thank you also to Inge for taking over testing for my experiments and to Birgit for coding all those clefts and questions for Chapter 4.

We are really lucky at the MPI to have excellent support. Maarten, thank you for your help with Presentation and Frinex. I don't know what the last chapter of this thesis would look like without you. Thank you Johan for dealing with all sorts of issues with experiment rooms. Thank you Alex for handling all my helpdesk questions and Paul and Ludy for patiently answering my many confused questions about data archiving. Thank you Karin and Meggie for always being kind and efficient and making Open Access extremely easy. Angela, thank for all your help with contract changes (real or imagined). Thank you also Marie-Luise for helping out with reimbursements. A big thank you also goes to Evelyn and Thy for handling all the paperwork I've produced over the years.

Kevin, thank you for making the IMPRS an awesome graduate school to be part of. I really enjoyed working with you on potentially a few too many events.

Thank you to the Open Science Community team over the years, Kristijan, Jeanette, Suhas, Leonore, Freddy, and Johannes. Brainstorming events, funding calls, and newsletters with you has been at times challenging, but mostly fun!

Thank you also to the rest of my Nijmegen friends for making my years there so special. Mischa, we weren't flatmates for as long as planned, but I really enjoyed sharing Chez Zijtaart (and Zijtaart) with you. Thank you for all the food and the long chats that were sometimes not about Belinda. Ashley, thank you for all the amazing dinners, be they turkeys, tacos, or the Body of Christ. Thank you Marieke for all the fun zoom borrels. Thank you also to Ellen for being my Lux buddy and all the long, occasionally depressing but always intriguing, conversations.

Thank you to all my friends around Europe and further afield. Tamara, thank you for all the lovely trips, for being my reminder of what the world

outside the ivory tower is like, and for being my knitting mentor. Thank you Candice for always being thoughtful and supportive and for making it all the way to Nijmegen to see me. Yoolim, thank you for all the postcards, cookies, and chats in recent months. Ένα μεγάλο ευχαριστώ επίσης τους φίλους μου στην Ελλάδα: Ιωάννα, Εβελίνα, Βούλα, Θοδωρή, Ελένη. Σ' ευχαριστώ Ευαγγελία που πάντα έβρισκες χρόνο να με δεις όταν ερχόμουν Ελλάδα και για τη βοήθειά σου με τη μετάφραση της περίληψης (και το μάθημα γραμματικής).

I am also really grateful to my family. Thank you Anna and Marcel for hosting me every time I came to Cologne. Thank you also for driving Elliot and me around Greece at a time when I really really needed a break. Μαμά, μπαμπά δεν καταλαβαίνατε πάντα τι ακριβώς κάνω στο διδακτορικό όμως είσασταν πάντα στο πλευρό μου. Ευχαριστώ.

Finally, thank you to my partner Elliot. Your support and encouragement have been invaluable. Thank you for letting me run by you basically my entire general introduction, general discussion, summary, and probably more. Meeting you was one of the best parts of my PhD.

Curriculum Vitae

Eirini Zormpa was born in Serres, Greece in 1991. She obtained a BA in English Language and Literature from the Aristotle University of Thessaloniki, where she first had a brush with psycholinguistic research, in 2013. After working as a teacher of English for a year, she decided she needed more research in her life. So, she moved to Reading, UK, to pursue an MSc in Language Sciences, where she worked on a range of topics on bilingualism. Eirini obtained her MSc in 2015, after which she worked as a research assistant at the University of Reading. In 2016, she moved to Nijmegen to pursue a PhD at the Psychology of Language department of the Max Planck Institute for Psycholinguistics. There she investigated how speaking and listening influence memory in conversations and also became fascinated with how to make research more transparent and reproducible. Eirini is now back in the UK, living in Cambridge.

Publications

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. doi: 10.1080/09658211.1510966

Zormpa, E., Meyer, A. S., & Brehm, L. E. (2019). Slow naming of pictures facilitates memory for their names. *Psychonomic Bulletin & Review 26*, 1675-1682. doi: 10.3758/s13423-019-01620-x

MPI Series in Psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda I. van Turennout*
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed class distinction in spoken-word recognition. *Alette Petra Haveman*
5. The acquisition of phonetic categories in young infants: A self organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk J.Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie H. van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*

16. Language-specific listening: The case of phonetic sequences. *Andrea Christine Weber*
17. Moving eyes and naming objects. *Femke Frederike van der Meulen*
18. Analogy in morphology: The selection of linking elements in Dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja Helena de Jong*
21. Fixed expressions and the production of idioms. *Simone Annegret Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Danielle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra Martine van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Henriëtte Wilhelmina Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl Jenny Judith Karin Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marijtje Elizabeth Debora Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Elisabeth Broersma*

35. Retrieving semantic and syntactic word properties: ERP studies on the time course in language comprehension. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren Batya Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure: Distinguishing contrastive from non-constrative variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel Haun*
42. The acquisition of auditory categories. *Martijn Bastiaan Goudbeek*
43. Affix reduction in spoken Dutch: Probabilistic effects in production and perception. *Mark Plumaekers*
44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Madalla Kooijman*
45. Space and iconicity in German sign language (DGS). *Pamela M. Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages: Convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Eva Wagner*
50. Lexical processing of morphologically complex words: An information theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo: A Papuan language of the Solomon Islands. *Claudia Ursula Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*
53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*

54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*

55. Processing the fine temporal structure of spoken words. *Eva Reinisch*

56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*

57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*

58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*

59. Lexical interactions in non-native speech comprehension: Evidence from electroencephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*

60. Processing casual speech in native and non-native language. *Annelie Tuinman*

61. Split intransitivity in Rotokas, a Papuan language of Bougainville. *Stuart Payton Robinson*

62. Evidentiality and intersubjectivity in Yurakaré: An interactional account. *Sonja Gipper*

63. The influence of information structure on language comprehension: A neurocognitive perspective. *Lin Wang*

64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*

65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*

66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*

67. The relevance of early word recognition: Insights from the infant brain. *Caroline Mary Magteld Junge*

68. Adjusting to different speakers: Extrinsic normalization in vowel perception. *Matthias Johannes Sjerps*

69. Structuring language: Contributions to the neurocognition of syntax. *Katrien Rachel Segaert*

70. Infants' appreciation of others' mental states in prelinguistic communication: A second person approach to mindreading. *Birgit Knudsen*

71. Gaze behavior in face-to-face interaction. *Federico Rossano*

72. Sign-spatiality in Kata Kolok: How a village sign language of Bali inscribes its signing space. *Connie de Vos*

73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*

74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt Witteman*

75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*

76. Territories of knowledge in Japanese conversation. *Kaoru Hayano*

77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder*

78. Contributions of executive control to individual differences in word production. *Zeshu Shao*

79. Hearing speech and seeing speech: Perceptual adjustments in auditory visual processing. *Patrick van der Zande*

80. High pitches and thick voices: The role of language in space-pitch associations. *Sarah Dolscheid*

81. Seeing what's next: Processing and anticipating language referring to objects. *Joost Rommers*

82. Mental representation and processing of reduced words in casual speech. *Iris Hanique*

83. The many ways listeners adapt to reductions in casual speech. *Katja Pöllmann*

84. Contrasting opposite polarity in Germanic and Romance languages: Verum Focus and affirmative particles in native speakers and advanced L2 learners. *Giuseppina Turco*

85. Morphological processing in younger and older people: Evidence for flexible dual-route access. *Jana Reifegerste*

86. Semantic and syntactic constraints on the production of subject-verb agreement. *Alma Veenstra*

87. The acquisition of morphophonological alternations across languages. *Helen Buckler*

88. The evolutionary dynamics of motion event encoding. *Annemarie Verkerk*

89. Rediscovering a forgotten language. *Jiyoun Choi*

90. The road to native listening: Language-general perception, language-specific input. *Sho Tsuji*
91. Infants' understanding of communication as participants and observers. *Gudmundur Bjarki Thorgrímsson*
92. Information structure in Avatime. *Saskia van Putten*
93. Switch reference in Whitesands. *Jeremy Hammond*
94. Machine learning for gesture recognition from videos. *Binyam Gebrekidan Gebre*
95. Acquisition of spatial language by signing and speaking children: A comparison of Turkish sign language (TID) and Turkish. *Beyza Sumer*
96. An ear for pitch: On the effects of experience and aptitude in processing pitch in language and music. *Salomi Savvatia Asaridou*
97. lncrementality and Flexibility in Sentence Production. *Maartje van de Velde*
98. Social learning dynamics in chimpanzees: Reflections on (nonhuman) animal culture. *Edwin van Leeuwen*
99. The request system in Italian interaction. *Giovanni Rossi*
100. Timing turns in conversation: A temporal preparation account. *Lilla Magyari*
101. Assessing birth language memory in young adoptees. *Wencui Zhou*
102. A social and neurobiological approach to pointing in speech and gesture. *David Peeters*
103. Investigating the genetic basis of reading and language skills. *Alessandro Gialluisi*
104. Conversation electrified: The electrophysiology of spoken speech act recognition. *Rósa Signý Gisladottir*
105. Modelling multimodal language processing. *Alastair Charles Smith*
106. Predicting language in different contexts: The nature and limits of mechanisms in anticipatory language processing. *Florian Hintz*
107. Situational variation in non-native communication. *Huib Kouwenhoven*
108. Sustained attention in language production. *Suzanne Jongman*
109. Acoustic reduction in spoken-word processing: Distributional, syntactic, morphosyntactic, and orthographic effects. *Malte Viebahn*

110. Nativeness, dominance, and the flexibility of listening to spoken language. *Laurence Bruggeman*

111. Semantic specificity of perception verbs in Maniq. *Ewelina Wnuk*

112. On the identification of FOXP2 gene enhancers and their role in brain development. *Martin Becker*

113. Events in language and thought: The case of serial verb constructions in Avatime. *Rebecca Defina*

114. Deciphering common and rare genetic effects on reading ability. *Amaia Carrión Castillo*

115. Music and language comprehension in the brain. *Richard Kunert*

116. Comprehending Comprehension: Insights from neuronal oscillations on the neuronal basis of language. *Nietzsche H.L. Lam*

117. The biology of variation in anatomical brain asymmetries. *Tulio Guadalupe*

118. Language processing in a conversation context. *Lotte Schoot*

119. Achieving mutual understanding in Argentine Sign Language. *Elizabeth Manrique*

120. Talking sense: The behavioural and neural correlates of sound symbolism. *Gwilym Lockwood*

121. Getting under your skin: The role of perspective and simulation of experience in narrative comprehension. *Franziska Hartung*

122. Sensorimotor experience in speech perception. *Will Schuerman*

123. Explorations of beta-band neural oscillations during language comprehension: Sentence processing and beyond. *Ashley Lewis*

124. Influences on the magnitude of syntactic priming. *Evelien Heyselaar*

125. Lapse organization in interaction. *Elliott Hoey*

126. The processing of reduced word pronunciation variants by natives and foreign language learners: Evidence from French casual speech. *Sophie Brand*

127. The neighbors will tell you what to expect: Effects of aging and predictability on language processing. *Cornelia Moers*

128. The role of voice and word order in incremental sentence processing. Studies on sentence production and comprehension in Tagalog and German. *Sebastian Sauppe*

129. Learning from the (un)expected: Age and individual differences in statistical learning and perceptual learning in speech. *Thordis Neger*

130. Mental representations of Dutch regular morphologically complex neologisms. *Laura de Vaan*

131. Speech production, perception, and input of simultaneous bilingual preschoolers: Evidence from voice onset time. *Antje Stoehr*

132. A holistic approach to understanding pre-history. *Vishnupriya Kolipakam*

133. Characterization of transcription factors in monogenic disorders of speech and language. *Sara Busquets Estruch*

134. Indirect request comprehension in different contexts. *Johanne Tromp*

135. Envisioning language - An exploration of perceptual processes in language comprehension. *Markus Ostarek*

136. Listening for the WHAT and the HOW: Older adults' processing of semantic and affective information in speech. *Juliane Kirsch*

137. Let the agents do the talking: On the influence of vocal tract anatomy on speech during ontogeny and glossogeny. *Rick Janssen*

138. Age and hearing loss effects on speech processing. *Xaver Koch*

139. Vocabulary knowledge and learning: Individual differences in adult native speakers. *Nina Mainz*

140. The face in face-to-face communication: Signals of understanding and non-understanding. *Paul Hömke*

141. Person reference and interaction in Umpila/Kuuku Ya'u narrative. *Clair Hill*

142. Beyond the language given: The neurobiological infrastructure for pragmatic inferencing. *Jana Bašnáková*

143. From Kawapanan to Shawi: Topics in language variation and change. *Luis Miguel Rojas Berscia*

144. On the oscillatory dynamics underlying speech-gesture integration in clear and adverse listening conditions. *Linda Drijvers*

145. Linguistic dual-tasking: Understanding temporal overlap between production and comprehension. *Amie Fairs*

146. The role of exemplars in speech comprehension. *Annika Nijveld*

147. A network of interacting proteins disrupted in language-related disorders. *Elliot Sollis*

148. Fast speech can sound slow: Effects of contextual speech rate on word recognition. *Merel Maslowski*

149. Reasons for every-day activities. *Julija Baranova*

150. Speech planning in dialogue - Psycholinguistic studies of the timing of turn taking. *Mathias Barthel*

151. The role of neural feedback in language unification: How awareness affects combinatorial processing. *Valeria Mongelli*

152. Exploring social biases in language processing. *Sara Iacozza*

153. Vocal learning in the pale spear-nosed bat, *Phyllostomus discolor*. *Ella Lattenkamp*

154. Effect of language contact on speech and gesture: The case of Turkish-Dutch bilinguals in the Netherlands. *Elif Zeynep Azar*

155. Language and society: How social pressures shape grammatical structure. *Limor Raviv*

156. The moment in between: Planning speech while listening. *Svetlana-Lito Gerakaki*

157. How speaking fast is like running: Modelling control of speaking rate. *Joe Rodd*

158. The power of context: How linguistic contextual information shapes brain dynamics during sentence processing. *Rene Terporten*

159. Neurobiological models of sentence processing. *Marvin Uhlmann*

160. Individual differences in syntactic knowledge and processing: The role of literacy experience. *Saoradh Favier*

161. Memory for speaking and listening. *Eirini Zormpa*