Influences on the magnitude of syntactic priming

Evelien Heyselaar

Max Planck Institute for Psycholinguistics

Series

Influences on the magnitude of syntactic priming

Evelien Heyselaar

Influences on the magnitude of syntactic priming

PhD Thesis, Max Planck Institute for Psycholinguistics

ISBN: 978-90-76203-86-7

Cover artwork by Evelien Heyselaar

Printed and bound by Ipskamp Drukkers bv

Het onderzoek in dit proefscrift werd mogelijk gemaakt door een beurs van de International Max Planck Research School for Language Sciences, gefinancierd door de Max Planck Society, Munchen, Duitsland.

Influences on the magnitude of syntactic priming

Proefschrift

ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

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

volgens besluit van het college van decanen

in het openbaar te verdedigen op vrijdag 2 juni 2017

om 10.30 uur precies

door

Evelien Suzanne Heijselaar

geboren op 23 augustus 1987

te Haarlem

Promotor: Prof. dr. P. Hagoort

Copromotor: Dr. K.R. Segaert (University of Birmingham, Verenigd Koninkrijk)

Manuscriptcommissie:

Prof. dr. A.P.A. Roelofs

Prof. dr. V.S. Ferreira (University of California – San Diego, Verenigde Staten)

Dr. F. Chang (University of Liverpool, Verenigd Koninkrijk)

Table of Contents

For my dad.

On February 14th 2013, you sent me the article *Neuroeconomics: Eyes, Brain, Business*, a Harvard Business school article about Looser and Wheatley's (2010) *The tipping point of animacy: How, when, and where we perceive life in a face* (Psychol Sci 21:1854–1862), the article that became the cornerstone of this thesis. Without your constant forwarding of interesting online articles, this thesis would have been a lot emptier.

I miss you.

1

General Introduction

And time and again it was described, that this or that wholly or partly empty syntactic schema preceded the actual formulation of an answer and in some way steered the effective speaking[1]

Karl Bühler (1934)

That words we just processed can influence our utterances is not a new phenomenon. Many people can relate to common mistakes such as "Get out of the clark" (where "car" was intended, said while glancing at a store front with the name "Clark's" printed on it; Harley, 1984) or "I am a sheep in lamb's clothing (where "wolf's clothing' was intended; Norman, 1981). These mistakes suggest that the words or sentences we have produced have been influenced by words or sentences in our surroundings. Apart from causing mismatch mistakes like the examples above, it can also manifest in the repetition of linguistic information.

Corpora have shown that people are prone to repeating the words they have just heard (Schenkein, 1980; Gries, 2005). In an attempt to experimentally measure this phenomenon, Levelt and Kelter (1982) conducted a question-and-answer telephone experiment in which they asked merchants either (a) "At what time does your store close?" or (b) "What time does your store close?" They observed that their interlocutors were much more likely to respond to (a) with a sentence also starting with a preposition, such as "At 6 o'clock", and to questions such as (b) with a noun-phrase, such as "6 o'clock." Although the authors interpreted this effect as proof that participants repeated words they just heard, this experiment also shows a repetition in grammatical structure, which is actually much more interesting.

Repeating the words in a previous sentence, also referred to as lexical priming, is explained by persistent activation (Levelt and Kelter, 1982). When we process a word, its mental representation becomes activated. This activation can last longer than the time it takes for us to comprehend or produce the word initially. Therefore, when the next utterance unfolds (whether it be comprehending or producing), the mental representation of the word is still partially activated. This makes it much more likely to be picked and used in the upcoming utterance, resulting in repeating words previously processed. However, grammar is much more abstract such that it very likely does not have a mental representation in the same sense that a word does. When it comes to building or understanding a sentence, there are constraints on how words can be placed in a sentence depending on their syntactic node (verb, subject, pronoun, etc.). This process of placing words in a sentence (or vice versa) has been termed unification (Jackendoff, 2002;

1 "Und immer wieder wurde dann beschreiben, dass dies oder jenes ganz oder teilweise leere syntaktische Schema der eigentlichen Formulierung einer Antwort vorherging und das faktische Sprechen irgendwie erkennbar steuerte."

Hagoort, 2005). As it is not very likely that we have representations of every possible grammatical structure which could be activated, the fact that syntactic structures get repeated between utterances suggests that it is the processes used for the unification of a previous sentence that are partially activated. Therefore, syntactic priming is a methodology with which to study the process of building sentences, not the sentences themselves. In other words: priming facilitates either access to (Pickering and Branigan, 1998) or the construction of syntactic structures (Chang et al., 2000, 2006; Jaeger and Snider, 2013) or a combination of both (Reitter et al., 2011).

In this thesis I will explore new characteristics of syntactic processing using syntactic priming as a methodological tool, and near the end try to explain the neurobiological basis of them. For the rest of this introduction I will briefly introduce concepts not mentioned in the introductions of the proceeding chapters.

Characteristics of Syntactic Priming

Although this repetition effect was an established phenomenon, Bock is recognized as being the first researcher to develop a paradigm to study this effect in a controlled and naturalistic manner at the syntactic level (Bock, 1986a). The task involved the experimenter reading a sentence aloud, which the participant then repeated. This “prime phase” was used so that the experimenter could control the exposure to different syntactic structures. The participant was then shown a picture and asked to describe it. This “target phase” was to see whether the participant would use the same structure they just produced in the prime phase in the subsequent utterance. This study is seen as more natural as it mimics an ongoing conversation more compared to the single question-and-answer procedure used by Levelt and Kelter (1982), and it additionally provides much more control on structure exposure compared to using corpora.

Since 1986, many psycholinguists have used and still use Bock’s design to investigate the characteristics associated with syntactic priming. The key characteristics can be narrowed down to: inverse frequency effects, cumulativity, lexical boost, and decay. I will briefly describe each in turn and how they contribute to our current view of syntactic processing.

Inverse frequency effects Less frequently occurring constructs prime more (Ferreira and Bock, 2006). For example, in Dutch and English the passive construct (i.e., “The man is being hugged by the woman”) is used much less frequently than the active construct (i.e., “The woman hugs the man”). Studies using these two constructs have shown that participants use passive structures more after a passive prime compared to after an active prime (Bock, 1986a; Bock et al., 1992; Hartsuiker and Kolk, 1998b; Bernolet et al., 2016). Of course, in an experiment, the exposure to passive and active structures are balanced, and hence many theoretical models suggest that this unexpected, increased used of uncommonly used structures is what drives the priming effect (Chang et al., 2000, 2006; Chang, 2002; Jaeger and Snider, 2013). Indeed, if exposure to the passive structure is trained such that it is not an uncommon anymore, the passive priming effect decreases (Segaert et al., 2011).

Cumulativity Studies have shown that repeated exposures to a prime increases the chances that the participant will use that construction in a future utterance (Chang et al., 2006; Jaeger and Snider, 2008; Chapter 2 this thesis). Hence there is growing evidence that priming is not necessarily a 0-back effect, but instead that participants continuously update their expectations and predictions of upcoming sentences such that they adapt their language behaviour throughout the entire experimental session. This element is used as evidence to support the learning view of priming.

Lexical boost The lexical boost effect refers to there being a boost in priming magnitude when words (nouns and/or verbs) are repeated between consecutive sentences. If there is intervening distractor material, this lexical boost effect is not seen (Hartsuiker et al., 2008). Hence it is thought that the lexical boost effect provides evidence that persistent activation does play a role in syntactic priming after all. After a sentence is processed, the words used remain active for a short period of time. If a following sentence repeats one or more of these words, the syntactic structure previously used (which is still partially active) is more easily re-activated compared to a structure that wasn't used. This therefore influences which structure the participant will use in their upcoming utterance. If other linguistic material is used instead, the partially activated words slowly lose their residual activation and hence can only influence consecutive sentences.

Decay As with the lexical boost, it has been shown that the prime can decay if material intervenes between prime and target. This has mostly been seen for written sentence production (Branigan et al., 1999; Wheeldon and Smith, 2003). Spoken language corpora have shown a decrease in priming magnitude depending on the time interval between prime and target (Gries, 2005; Reitter, 2008) although the current consensus is that written priming decays, whereas spoken priming persists over longer instances (Hartsuiker and Kolk, 1998b; Bock and Griffin, 2000; Branigan et al., 2000; Hartsuiker et al., 2008). This apparent disconnect is elaborated on further below.

The Role of Memory in Syntactic Priming

It seems that certain characteristics have contradictory descriptions. For example, cumulativity describes that the chances of repeating a structure increases over the length of an experimental session, yet the decay suggests the opposite: there is a decrease in priming effect with intervening material between prime and target. This has stimulated the idea that perhaps the behavioural measurement of priming is actually two different things: 1) an implicit learning mechanism that causes consistent and persistent syntactic priming, and 2) the explicit memory of the prime sentence's surface structure (i.e., individual words used) that causes a short-lived effect (Bock and Griffin, 2000; Chang et al., 2006; Ferreira and Bock, 2006). Indeed, when one controls for lexical overlap, there is still syntactic priming, but without the short-decay (Hartsuiker et al., 2008; Bernolet et al., 2016).

Memory tasks have been used to assess the explicit memory contribution of immediate prime-target pairs and have shown that participants do have a good memory for these trials (Bernolet et al., 2016). However, the implicit memory of syntactic priming is harder to test. Empirical demonstrations of priming studies suffer from the problem of possible explicit contamination. That is, in healthy participants it is difficult to rule out the possibility that priming effects may be mediated by explicit memory processes as well. Therefore, for **Chapter 2** we investigate the role of implicit memory by testing patients with Korsakov's amnesia on a syntactic priming task. These patients display deficits in all subdomains of explicit memory, but implicit learning remains intact, making them an ideal patient group to investigate the role of implicit memory in syntactic priming.

The Role of Social Influences in Syntactic Priming

In addition to implicit and explicit memory underlying and thereby influencing the efficiency of syntactic processing, such that "using procedures that are already activated may ease the demands of message formulation and actually contribute to fluency" (Bock, 1986; pg. 379), the opinion a participant has of their interlocutor seems to influence the magnitude of syntactic priming as well (Balcetis and Dale, 2005; Weatherholtz et al., 2014). This is similar to what has been seen in the field of behavioural priming.

Behavioural priming refers to the phenomenon in which people copy the behaviour of their partner. This can be body posture, foot tapping, face scratching, and many more (Chartrand and Bargh, 1999; Hess et al., 1999). Many researchers in the field of behaviour priming suggest that we mimic our partners to indicate that we like them; that mimicry is an unconscious social cue (Lott and Lott, 1961). Could language mimicry be cut from the same cloth and also function as an unconscious social cue?

The one problem with investigating social influences is that it is a very abstract, very personal thing. One never knows precisely what characteristics make one person seem more likeable compared to the next. In priming studies, confederates are usually used as the interlocutor to ensure an equal exposure to the grammatical structures tested. However, the role of the confederate has recently been scrutinized, in terms of the potential for artificially produced signals that may influence the behaviour exhibited by the participant (Kuhlen and Brennan, 2013).

We suggest a methodology to overcome this obstacle: we have participants interact with digital partners ("avatars") in Virtual Reality. This will minimize noise in terms of social opinion ratings as the computer will behave exactly the same with every participant. In **Chapter 3** we answer the question, is this ecologically valid? Does interacting with a computer accurately reflect interaction with a confederate, specifically looking at syntactic priming. In **Chapter 4** we use our established methodology to investigate whether social opinion can influence the magnitude of the structural priming effect, to determine whether language repetition also reflects social opinion similar to those seen in the behavioural mimicry literature.

The Role of Attention in Syntactic Priming

The influence of social opinion on syntactic priming is interesting, but it is also difficult to explain. Social opinion is not one area in the brain that modulates the strength of a behaviour, rather it is likely a combination of perceptual and cognitive operations. Therefore, in this last portion of the thesis I propose one piece of this puzzle that could influence behaviour in response to the opinion one has of their interlocutor: distractibility.

In **Chapter 5** we use a dual-task to determine if attention is even necessary for syntactic processing. If it isn't, then it cannot be an underlying feature of the umbrella term that is social opinion. In **Chapter 6** we provide neuroimaging evidence in the form of alpha oscillations that opinion is driven (at least partially) by attention. In the final chapter of this thesis, I bring together my work so far, to show what we have uncovered not only about the relationship of social opinion and syntactic priming, but also how this can be explained from a neuroscientific perspective. How does our brain differ when we interact with people we like compared to people we don't like?

2

The role of procedural memory in the skill for language: Evidence from syntactic priming in patients with amnesia

2

Syntactic priming, the phenomenon in which participants adopt the linguistic behaviour of their partner, is widely used in psycholinguistics to investigate syntactic operations. Although the phenomenon of syntactic priming is well documented, the memory system that supports the retention of this syntactic information long enough to influence future utterances, is not as widely investigated. We aim to shed light on this issue by assessing patients with Korsakoff's amnesia on an active-passive syntactic priming task and compare their performance to controls matched in age, education, and premorbid intelligence. Patients with Korsakoff's syndrome display deficits in all subdomains of explicit memory, yet their implicit memory remains intact, making them an ideal patient group to determine which memory system supports syntactic priming. In line with the hypothesis that syntactic priming relies on procedural memory, the patient group shows strong priming tendencies (12.6% passive structure repetition). Our healthy control group did not show a priming tendency, presumably due to cognitive interference between declarative and non-declarative memory. We discuss the results in relation to amnesia, aging and compensatory mechanisms.

Adapted from: Heyselaar E, Segaert K, Walvoort SJW, Kessels RPC, Hagoort P (under review) The role of non-declarative memory in the skill for language: Evidence from syntactic priming in patients with amnesia *Neuropsychologia.*

Introduction

The human language system is often characterized by a tripartite architecture (Jackendoff, 2002) that enables us to map sound onto meaning (in listening) or meaning onto sound (in speaking). Next to sound and meaning, there is syntax, which enables the well-formed grouping of words into longer utterances. At a very general level, for all three information types (sound, syntax, meaning), one can make a distinction between two crucial components. The one relates to the common assumption that the basic building blocks of linguistic knowledge get encoded and consolidated in the course of language acquisition. This is what we refer to as the *Memory* component of the human language system, and is more usually called the mental lexicon in the field of psycholinguistics. Crucially, however, language processing is more than the retrieval of lexical knowledge and goes beyond the simple concatenation of retrieved lexical items. The expressive power of human language derives from the possibility to combine elements from memory in often novel ways. This creative aspect led Wilhelm von Humboldt (1829) to characterize language as a system which *"makes infinite use of finite means."* This process of deriving new and complex meaning from the lexical building blocks is referred to by some as *Unification* (Hagoort, 2005, 2013, 2016). This process supports the on-line assembly of lexical building blocks into larger structures, with contributions from context and general world knowledge. It instantiates what in linguistic theories is often called the compositionality of language. Although the mental lexicon is part of semantic memory, and hence a component of declarative memory (Ullman, 2001; Hagoort, 2005), it is less clear which memory structure supports the on-line assembly of utterances that are not prestored in the mental lexicon. It has been argued (Ullman, 2001) that the on-line composition (speaking) or decomposition (listening/reading) of sound, morphological and syntactic structures is subserved by procedural memory (Gupta and Cohen, 2002). Here we investigate a group of patients with severe amnesia that might provide relevant information on the contribution of procedural memory to human language skills, more in particular to the Unification component of the language system.

A core process in language production and comprehension is the production and comprehension of the syntactic relations between the lexical items in an utterance; i.e., the processing of the relationships between words in a sentence. The same words can be combined, but in different syntactic roles (e.g., subject, object), to produce different meanings (*the man kisses the woman/the woman kisses the man*) or different words fulfilling the same syntactic roles can be combined to produce the same meaning (*the man kisses the woman/the woman is kissed by the man*). Without a functioning syntactic processing system, the ability to understand language as well as to produce it is severely impaired. As language production and comprehension are so tightly linked, interlocutors are prone to repeat the syntactic structure in which their partner formulates her utterance. Indeed, corpus studies have shown that interlocutors adapt their syntactic language behaviour to match that of their partner (Giles and Powesland, 1975).

In 1982, Levelt and Kelter were the first to experimentally reproduce this repetition of syntax; they showed that the question "*On which instrument does Paul play?*" was more commonly answered (89%) with "*On the piano*" as opposed to "*The piano*" by the 36 participants they tested. The language adaptation behaviour, referred to in this article as syntactic *priming*, but also known as *accommodation* or *alignment,* has been used in a wide-range of applications. Syntactic priming studies have shown that abstract linguistic structures have a basis in psychological reality (Bock, 1986a), how these are acquired during language development (Kidd, 2012)**,** and which role syntactic priming plays in social cueing (Balcetis and Dale, 2005). However, the memory system that is needed to retain this linguistic information long enough to be used in producing utterances has not been seriously investigated.

Most studies that have examined the retention of linguistic information over time did not distinguish between different memory types. However, studies that investigated the effect of intervening irrelevant linguistic information or just time itself (Branigan et al., 2000), using either spoken (Bock and Griffin, 2000) or written modalities (Hartsuiker et al., 2008), did not observe a significant decrease in priming ability. Although the primed structures may remain active over some intervening trials, the length of the decay (sometimes even a week; Kaschak, Kutta, & Schatschneider, 2011) does not rule out that the participant may have consciously learnt the relevant linguistic structures. This points towards the involvement of declarative memory, the memory that underlies the acquisition, representation, and use of knowledge about facts and events.

At the same time, other studies have suggested that priming might be a form of statistical learning that is non-declarative and implicit: participants automatically and unconsciously pick up on the frequency of event occurrences, which could explain why they produce these events with increasing probability over the length of the experimental session (Bock and Griffin, 2000; Kaschak et al., 2011a; Jaeger and Snider, 2013). Indeed a detailed computational model has been developed which supports these claims (Chang et al., 2000, 2006; Chang, 2002). One critical note is that empirical demonstrations of priming studies suffer from the problem of possible explicit contamination. That is, in healthy participants it is difficult to rule out the possibility that priming effects may be mediated by explicit memory processes as well.

The most direct method to ensure that there is no influence of the declarative memory system is to measure participants that have amnesia. Until now, only one study has used this approach: Ferreira and colleagues (2008) had patients with explicit memory deficits complete a syntactic priming task (a task that focuses only on grammatical adaptation in language behaviour) and compared their performance to age- and IQ-matched controls. Their results showed that patients' ability to repeat syntactic structures did not differ significantly from the control group, even though their explicit memory performance was significantly worse compared to the controls. This led the authors to conclude that syntactic priming does not require explicit memory. However, this is only a single study, which examined only four patients with a mixed aetiology. The mixed aetiology could potentially be a confound, as both the explicit and implicit memory systems have extensive neural networks, and thus lesions in different areas may not affect the four patients to the same extent.

The explicit, or declarative, memory system is mainly based in the diencephalon and the medial temporal lobe (MTL) structures. These include the hippocampus proper, the entorhinal cortex, the perirhinal cortex, and the parahippocampal cortex (Squire and Knowlton, 2000; Suzuki and Eichenbaum, 2000; Squire and Dede, 2015). The hippocampus projects to the midline diencephalic nuclei, including the mammillary bodies and portions of the thalamus (Kopelman, 2014). This diencephalic-MTL circuitry is involved in several memory related functions, including encoding, consolidation, and retrieval of new memories (Squire and Knowlton, 2000; Eichenbaum and Cohen, 2001), although memories eventually become mostly independent of the medial temporal lobe structures and dependent upon neocortical regions, particularly the temporal lobes (Hodges and Patterson, 1997; Squire et al., 2001). For language, memory for items stored in the mental lexicon has usually been related to inferior, middle, and superior temporal lobe regions (Hagoort, 2013, 2014; Hagoort and Indefrey, 2014; Ullman, 2001).

The implicit, or non-declarative (procedural), memory system is composed of an extensive neural network, with the root in the frontal-striatal circuits and branching out to include portions of the parietal cortex, superior temporal cortex, and the cerebellum (De Renzi, 1989; Schacter and Tulving, 1994, Squire and Dede, 2015). The input to the basal ganglia (including the striatum) depend upon the type of information involved; for example motor learning might be projected from the supplementary motor area (SMA) and pre-supplementary motor area (Middleton and Strick, 2000), whereas syntax-related combinatorial operations (i.e., syntactic unification) could be projected from areas such as Broca's region (Conway and Christiansen, 2001). The information is then processed in the basal ganglia and projected back to prefrontal cortex, closing the loop. As the network is so extensive, it is imperative to ensure that whatever the cause of the patient's explicit memory deficit, implicit learning is not affected.
Korsakoff's syndrome is a neurological disorder caused by a chronic deficiency of thiamine (vitamin B_1) due to severe malnutrition usually associated with chronic alcoholism. Patients display profound amnesia due to bilateral lesions to the thalamus and mammillary bodies (Pitel et al., 2014) which, as mentioned above, are structures relevant for the encoding and consolidation of new memories via the explicit/declarative memory system. Patients therefore display deficits in all subdomains of explicit memory, but implicit learning remains intact (Cermak et al., 1991; Oudman et al., 2011), making them an ideal patient group to include

in this study.

In this study we aim to shed light on which memory system underlies syntactic priming. To control for any influence of the explicit memory system, we will be comparing the performance of amnesia patients with age-, education-, and premorbid intelligence-matched controls in a syntactic priming task. Overall, if syntactic priming is supported by non-declarative/procedural memory, the amnesia patients should show robust priming effects.

Materials and methods

Participants

Patients with amnesia Eighteen patients with Korsakoff's syndrome (13 men) were recruited from the Centre of Excellence for Korsakoff and Alcohol-Related Cognitive Disorders at the Vincent van Gogh Institute of Psychiatry in Venray, The Netherlands. For all patients, the current intelligence level of each participant had to be in concordance with the estimation of pre-morbid functioning based on occupational and educational history, to exclude possible alcohol-related dementia (Oslin et al., 1998). All patients fulfilled the DSM-5 criteria for Alcohol-Induced Major Neurocognitive Disorder, Amnestic Confabulatory Type (American Psychiatric Association, 2013), and the criteria for Korsakoff's syndrome described by Kopelman (2002). The diagnosis was supported by extensive neuropsychological testing. All patients were in the chronic, amnestic stage of the syndrome. None of the patients were in the confusional Wernicke psychosis at the moment of testing. No brain abnormalities were found that are at odds with the diagnosis of Korsakoff's syndrome (i.e., stroke, tumour, etc.). Patients had an extensive history of alcoholism and nutritional depletion, notably thiamine deficiency, verified through medical charts or family reports.

All testing occurred after the patients had been abstinent from alcohol for at least six weeks. The study was approved by the Vincent van Gogh Institutional Review Board (*Commissie Wetenschappelijk Onderzoek Participatie* U14.012). All the patients were informed about the study by the clinical staff and asked whether they were willing to participate; if so, a written informed consent was obtained.

Table 2.1 Descriptive statistics of the two participant groups. Education level was measured using seven categories in accordance with the Dutch educational system (1 = less than primary school; 7 = academic degree; Verhage, 1964); premorbid intelligence (IQ) was measured using the Dutch version of the National Adult Reading Test (NART; Schmand, Lindeboom, and van Jarskamp, 1992). There are no significant differences between groups (Mann-Whitney U test, $p > .077$)

		Controls	Amnesia patients	*p* value
Age	(mean(SD))	62.0 (6.73)	62.2 (8.0)	.919
Education level	(mode(range))	5 (2)	4 (6)	.077
NART-IQ	(mean(SD))	99.3 (20.78)	95.50 (20.1)	.451

Controls Eighteen healthy participants (8 men) were recruited from the Max Planck Institute for Psycholinguistics database and tested at the institute. These participants were matched with the patients in age, education level, and verbal IQ (see Table 2.1). No control participants reported any neurological deficits or psyczhiatric disorders and none had been treated for addiction. At the time of testing, none of the patients were taking any psychoactive medication. The study was approved by the ethics commission of the Faculty of Social Sciences at Radboud University, Nijmegen (Ethics Approval # ECG2013-1308-120).

Materials

All participants completed a syntactic priming experiment and were also tested on their implicit and explicit memory ability. For the Korsakoff patients, the implicit and explicit memory test scores were obtained as part of the routine neuropsychological assessment. The healthy controls completed the syntactic priming task, implicit memory task, and explicit memory task (in that order), in one session of approximately one and a half hours. Syntactic priming data for all participants (patients and controls) were collected by the same experimenter (E.H.).

Implicit memory test For healthy controls, it is near to impossible to measure pure implicit memory without possible explicit memory contamination, as outlined previously. Therefore, this test was mainly executed to ensure that the *patients* still had implicit learning ability.

To test implicit memory, we used the Fragmented Pictures Test (Kessels et al., 2011). Participants are shown a set of 7 line drawings, in a sequence of 8 pictures of decreasing degradation. Each picture in the sequence was presented for 3 seconds. The participant is instructed to name the picture, to answer only if they are quite sure and not to guess. For each line drawing sequence, the sequence number is recorded at which the participant correctly identified the picture. There are 3 consecutive runs of this task and a fourth run after a delay of 10 minutes. The participant's performance reflects their average sequence number out of the 8 pictures, for each trial type (3 learning trials and one delay trial).

Explicit memory yest To test explicit memory, we used the Rivermead Behavioural Memory Test - Third Edition (RBMT-3; Wester, Van Herten, Egger, and Kessels, 2013; Wilson et al., 2008). This extensive test battery consists of a range of everyday memory types (face recognition, picture recognition, story recall, prospective memory route recall, etc.). The participant's overall performance (Global Memory Index; GMI) is a summary of their scores at each subtest, corrected for age.

Syntactic priming To test syntactic priming ability, we presented 80 prime-target picture pairs.

Stimulus pictures The pictures used in this task have been described elsewhere (Segaert et al., 2011). The stimulus pictures depicted 40 transitive events such as *kissing, helping* or *strangling* with a depiction of the agent and patient of this action. Each event was depicted with two pairs of adults and two pairs of children. One male and one female actor were shown in each picture, and each event was depicted with each of the two actors serving as the agent. To prevent the forming of strategies, the position of the agent (left or right) was randomized. Studies have suggested that lexical repetition (a boost in priming magnitude seen when verbs or nouns are repeated in consecutively presented stimuli) is based on explicit memory (Hartsuiker et al., 2008; Kidd, 2012; Bernolet et al., 2016). Thus, to ensure that the control group did not have an advantage over the patient group, no verb or actor type (adult/children) was consecutively repeated. Studies have shown that priming still occurs without this lexical repetition (Branigan et al., 1999; Hartsuiker et al., 2008; Bernolet et al., 2016).

Each transitive picture had three versions: one grayscale version and two colour-coded versions with a green and a red actor (which elicited sentences with either an active or passive transitive). Fillers elicited intransitive sentences, depicting events such as *running, singing, bowing* with one actor (in greyscale or green).

Experimental Procedure

Participants were instructed to describe pictures with one sentence, naming the green actor before the red actor if the actors are depicted in colour. This allowed us to manipulate whether the prime sentence produced had an active or a passive syntactic structure. Figure 2.1 depicts the order of events for the syntactic priming task. If the actors were not depicted in colour, the participants could describe the photo however they wished, producing voluntarily either an active or passive sentence. To ensure that the patients did not forget the instructions, they were written at the top of the screen for each picture. Additionally, the verb that the picture is depicting was written at the bottom of the screen.

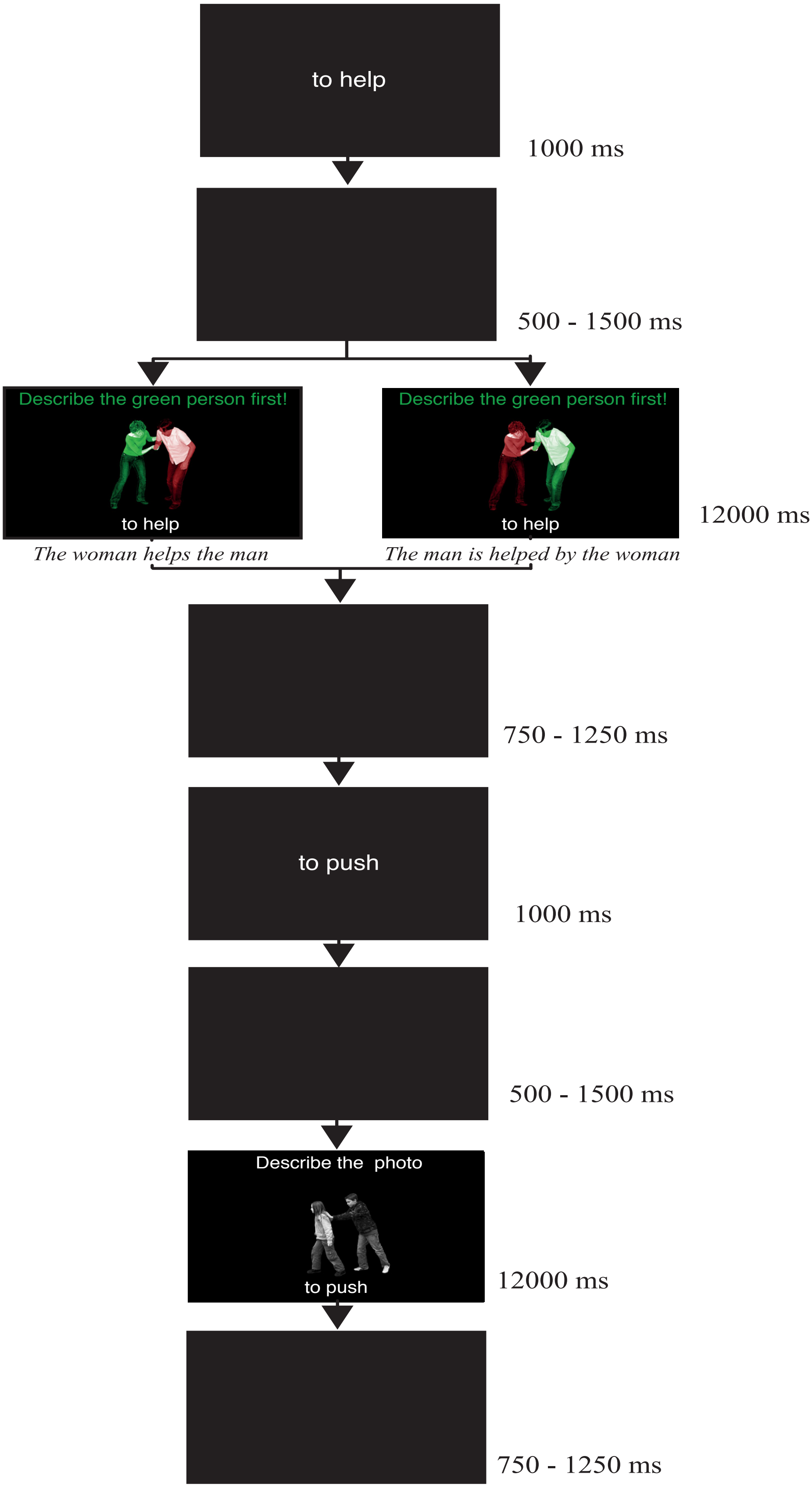

Figure 2.1 Order of events for the syntactic priming task. Pictures are presented until a response is produced.

Each trial consisted of a prime (a coloured picture) followed by a target (a greyscale picture). There were 20 passive prime trials (a passive picture followed by a transitive greyscale target), 20 active prime trials (an active picture followed by a transitive greyscale target), and 20 baseline trials (an intransitive picture followed by a transitive greyscale target), all randomized in one experimental session. This resulted in 80 transitive pictures and 20 intransitive pictures. The baseline trials allowed us to measure the frequency of producing active and passive transitives on subsequent targets without any immediate prior influence.

All pictures were presented until the participant responded. Filler trials were also included (20% of all trials, consisting of an intransitive prime followed by an intransitive target). This brings the total up to 60 intransitive pictures and 100 transitive pictures.

Coding and Analysis

Responses during the syntactic priming task were manually coded by the experimenter as either active (0) or passive (1). An independent rater blind to the purpose of the experiment verified that the coding was performed correctly. Trials in which the descriptions did not match one of the coded structures were discarded. Target responses were included in the analysis only if 1) both actors and the verb were named (a sentence naming only one of the actors does not qualify as a transitive sentence) and 2) the structures used were active or passive. In total 127 trials (9.34%) in the patient group were discarded; 144 trials (8.38%) in the control group were discarded. One patient had over 30% unusable trials and was discarded entirely from the data set; that patient's age-, education- and IQ-matched control was also discarded to maintain an equal number in each group.

The responses were analyzed using a mixed-effects logit model, using the glmer function of the lme4 package (version 1.1.-4; Bates, Maechler, and Bolker, 2012) (Looser & Wheatley, 2010; Tinwell, Grimshaw, & Williams, 2010) in R (R Core Development Team, 2011). Target responses were coded as 0 for actives and 1 for passives in factor *Prime*. We used a maximal random-effects structure (Jaeger, 2009; Barr et al., 2013): the repeated-measures nature of the data was modelled by including a per-participant and per-item random adjustment to the fixed intercept ("random intercept"). We began with a full model and then performed a step-wise "best-path" reduction procedure, removing interactions before main effects, to locate the simplest model that did not differ significantly from the full model in terms of variance explained. Factorial predictors were dummy coded (all means compared to a reference group) and all numeric predictors were centered. We included the factor *Cumulative Passive Proportion* to reflect any learning trend exhibited by the participants. This factor was calculated as the proportion of passives out of the total transitive responses produced on the target trials before the current target trial.

We used intransitives as the reference group for *Prime*. Collinearity between factors was low (VIF < 1.37).

Results

Figure 2.2 shows the results for the explicit and implicit memory tests. Controls showed a significantly higher explicit memory result (M = 96.18, SD = 19.80) compared to the amnesia group (M = 62, SD = 5.38, Mann-Whitney $U = 0$, $p < .001$), who performed in the impaired range (in the 20th percentile).

In the amnesia group, a significant learning curve was present for the Fragmented Picture Task performance[2] (Friedman $\chi^2(3) = 39.686$; $p < .001$), with an increase in performance on trial 3 compared to trial 1 (Friedman $\chi^2(3) = -3.298$; $p < .001$), and an increase in performance between trial 3 and the delayed trial (Friedman $\chi^2(3) = -3.236$; $p < .001$) indicating that the patients retained information between the trials, even with a 10-minute delay. As the amnesia patients performed within the impaired range on their explicit memory test, presumably their performance in the implicit memory task relies mostly on their non-declarative/procedural memory system.

Figure 2.3 summarizes the relative proportion of passive target responses after each prime structure. The fixed effects of the model fit for these data are summarized in Table 2.2.

Table 2.2 Summary of the best mixed logit model for passive vs. active response choices.

Results for the Control Group

Predictor	coefficient	*SE*	*Wald Z*	*p*	
Intercept (baseline)	-2.84	0.35	-8.07	< .001	***
Active Prime	-0.59	0.55	-1.07	> .250	
Passive Prime	0.41	0.42	0.97	> .250	
Cum. Passive Prop	8.17	1.27	6.41	< .001	***

N = 927, log-likelihood = -246.9

Results for the Amnesia Group

Predictor	coefficient	*SE*	*Wald Z*	*p*	
Intercept (baseline)	-2.20	0.26	-8.30	< .001	***
Active Prime	-0.23	0.42	-0.55	> .250	
Passive Prime	0.90	0.34	2.64	.008	**
Cum. Passive Prop	8.35	0.78	10.68	< .001	***

N = 909, log-likelihood = -324.0

2 Note that only 14 out of 17 amnesia patients gave consent to conduct the implicit memory task.

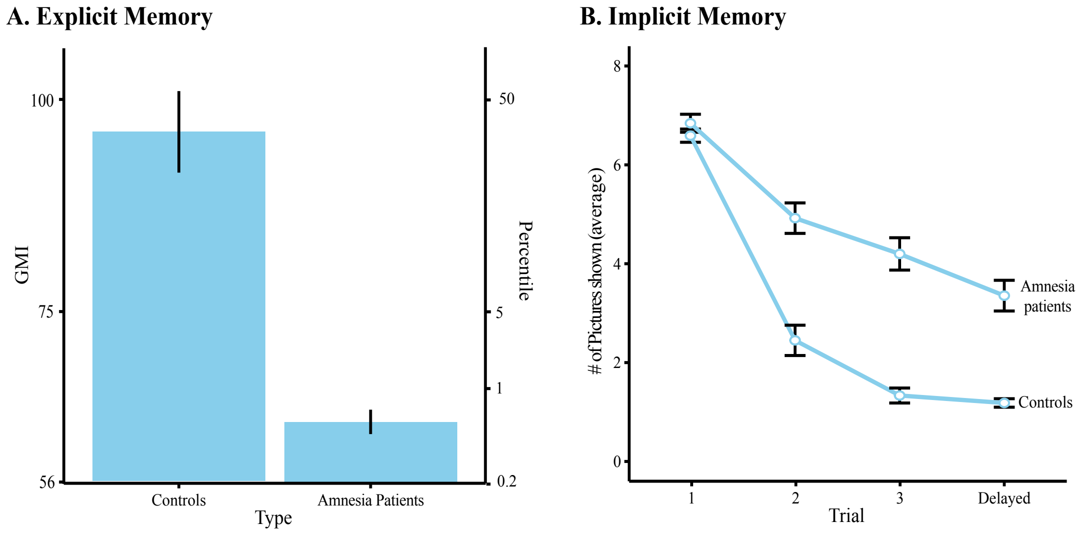

Figure 2.2 Results of memory tests. Only 14 (out of 17) amnesic patients agreed to complete these tests. **A. Explicit memory**: Controls showed significantly higher explicit memory performance compared to the amnesic group on the RBMT-3 ($p < .001$). **B. Implicit memory**: Amnesia patients showed a significant learning trend on the Fragmented Pictures Test, indicating that their non-declarative memory ability is still intact. Error bars represent standard errors of the mean.

The negative estimate for the intercept indicates that in the baseline condition active responses were more frequent than passive responses. For both groups, there was no increase in active responses following active primes, compared to baseline ($p = .283$).

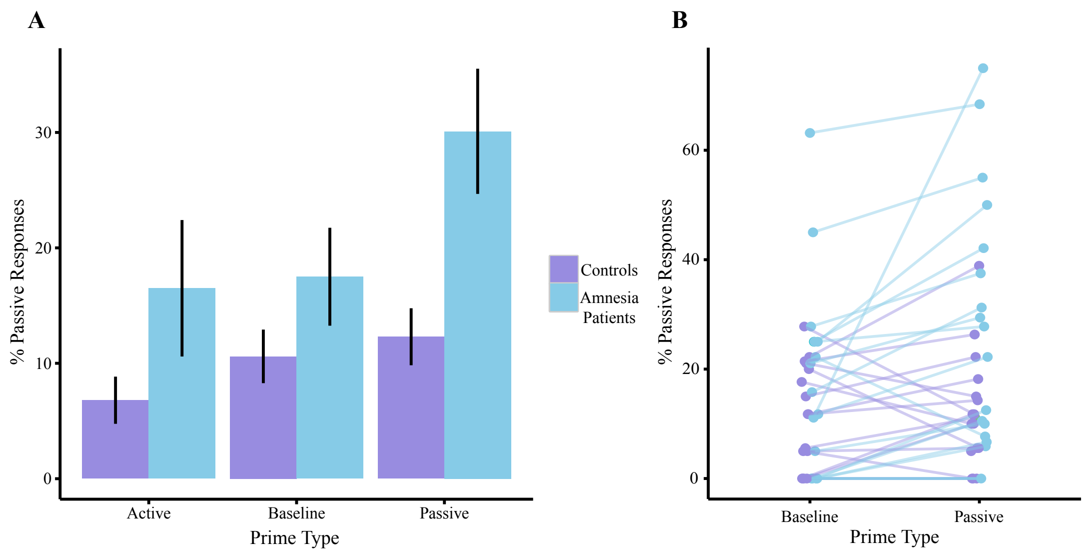

Figure 2.3 Percentage of passive responses per prime per group. Following a passive prime, the production of a passive sentence increases with 12.6% for the amnesia group and with 1.7% for the control group compared to the baseline condition. In line with previous research, there were no priming effects for actives. Panel **A** shows the average of both groups (error bars represent standard error), whereas panel **B** plots the individual performances for the baseline and passive prime trials.

Significantly more passives were produced compared to baseline (*p* = .008) by the amnesia patients, indicating that despite their explicit memory deficits, they were still able to retain syntactic information. This is also reflected in their *Cumulative Passive Proportion*, which was calculated as the proportion of passives out of the total transitive responses produced on the target trials before the current target trial. Any passives produced during prime trials are not included in this calculation. A positive and significant *Cumulative Passive Proportion* therefore suggests that the proportion of passives previously produced positively influences the probability of producing a passive on the current target trial. In other words, there is a cumulative effect of syntactic priming (i.e., the more passives produced, the stronger the effect), supporting a learning effect of priming. This factor also reflects any delayed priming influence, as opposed to the *Prime* condition, which reflects immediate prior influence. As the patients have an impaired explicit memory system, this ability is most likely supported by their implicit memory ability.

Somewhat unexpectedly, the control group did not show a significant priming effect. They demonstrated a significant learning trend, as reflected in *Cumulative Passive Proportion*, suggesting they accumulated information, but not enough to produce higher than a 1.7% priming effect.

We included *Education Level* as a factor in the full model, as the difference between the groups was nearly significant (*p* = .077). Including this factor did not make the fit of the model significantly better (*p* > .290), and therefore this was not included in the best model reported above.

Discussion

Our results support the theory that syntactic priming is based on implicit, or non-declarative/procedural, memory. We examined 17 patients with amnesia due to Korsakoff's syndrome in a syntactic priming experiment. Memory tests supported the claim that these patients did have a severely impaired explicit memory system, yet a functional implicit one. Fully in line with predictions, the Korsakoff patients showed a strong passive priming tendency, providing unequivocal support that syntactic priming is an implicit memory process. Somewhat unexpectedly, however, our healthy control group did not show a significant priming effect.

Our results are at odds with an earlier study investigating the same issue. Ferreira and colleagues (2008) tested four amnesia patients with a mixed aetiliogy with four age-, education-, and IQ-matched healthy controls. They found a significant priming effect not only for their amnesia group, but also for their control group. Although the ages of their participants are younger than ours (*M:* 50.875 vs. 62.09 years), a point we will address below, another major difference between our study and that of Ferreira and colleagues are the syntactic structures used (dative vs. transitive). Research has suggested that priming effects for transitives are generally weaker and more fragile than priming effects for datives (Bock and Griffin, 2000) even though the characteristics of the priming effects are comparable (Bernolet et al., 2016). This may be one explanation as to why our results differ in terms of the control group, although other potential explanations are addressed below.

In this study we used a production-production design, in which the participant's picture description of the colour coded pictures would act as his or her own prime when describing the grey pictures. However, other methods exist to test syntactic priming ability. These designs include listening to the prime being described by either a recording or a confederate (comprehension design; Bock and Griffin, 2000; Menenti, Gierhan, Segaert, and Hagoort, 2011), or having the participant read the prime sentence and then write out the target sentence (Branigan et al., 1999; Hartsuiker et al., 2008), or any combination of the above. As all of these have shown robust syntactic priming effects, it suggests that the underlying mechanism should be independent of the modality used, and therefore we are confident that our results are applicable to other modalities of priming as well and are not unique to the production-production methodology.

Of course, if the underlying mechanism is independent of the modality used, then investigating which brain regions are involved in all modality types should help elucidate the core of syntactic operations, and thereby which memory type(s) supports it. A neuroimaging study by Segaert, Menenti, Weber, Petersson, and Hagoort (2012) did just that: they compared the brain areas involved when syntactic priming was measured in a sentence production task with measurements in a sentence comprehension task. They demonstrated that in both cases adaptation effects were found in the left inferior frontal gyrus (IFG), left middle temporal area (MTG), and bilateral supplementary motor area (SMA). These three areas are known to be involved in language processing, in particular in the unification of language information (IFG; Hagoort, 2003, 2005; Snijders et al., 2009), in the process of sequencing syllable structures (SMA; Segaert et al., 2012), and in the retrieval of lexical-syntactic information from memory (MTG; Snijders et al., 2009), respectively. This latter process is thought to refer to the retrieval of syntactic frame information (Vosse and Kempen, 2000) but the memory system underlying this storage and retrieval is still unknown, and hence no clear conclusions could be drawn about which memory system underlies syntactic priming from this study.

In the current study, we extended these findings by using a lesion model. By examining patients with a deficit in a specific cognitive system, in this case explicit memory, we can determine whether that system is involved in the behaviour of interest. In our study, we show that amnesia patients with explicit memory deficits are still able to show robust syntactic priming ability, further supporting the claim that syntactic priming is supported by procedural memory. This result is unexpected given the language impairments typically seen in other types of patients. For instance, patients with primary progressive aphasia have a strong deficit in single word comprehension (Mesulam et al., 2015), due to cortical atrophy in the left anterior temporal lobe. This area is part of the declarative knowledge base for lexical items. In contrast patients with Parkinson's disease are impaired in producing correct inflections when these are regular and hence rule-based (Ullman et al., 1997; Ullman, 2001). The degeneration of the basal ganglia in these patients affects the procedural memory system. These patients show a deficit in procedural aspects of word formation, but not in retrieval of lexical information from memory. The Korsakoff patients in our study show, on the other hand, a preservation of implicit knowledge relevant for syntactic encoding, i.e., the formation of grammatically well-formed sentences, while at the same time suffering from serious impairments in declarative memory. All in all, this is a strong indication that language processing recruits multiple memory systems, at least encompassing declarative and procedural memory. Within the language domain, syntactic processing is a way to solve the problem of serial order in speaking (Lashley, 1951); that is, to put the lexical items retrieved from declarative memory (the mental lexicon) in a specific word order. In general terms, procedural memory is known to be involved in sequencing and timing (Nemeth et al., 2011; Dehaene et al., 2015). This might explain why procedural memory is centrally involved in syntactic skills.

In remarkable contrast to the patients and to the younger healthy participants (Segaert et al., 2011), the age- and education-matched controls failed to show a syntactic priming effect. The finding that an older control sample exhibits less priming compared to the cognitively impaired patients is not a new observation: in a study testing patients with Broca's aphasia, the patient group showed stronger syntactic priming effects compared to the healthy controls (Hartsuiker and Kolk, 1998a).

The one consistent element between the Hartsuiker and Kolk (1998a) study and our study that sets them apart from other syntactic priming studies is the age of the participants. So far, most syntactic priming studies are limited to using the undergraduate population: students around 20 years of age. As patients with general amnesia, Korsakoff's syndrome or Broca's aphasia are on average older and may also have an average or below-average education, most patient studies use non-academic older healthy controls. Therefore, the somewhat unexpected lack of a priming effect seen in the control group could be due to age or other confounding factors. We will discuss two possible explanations for the lack of a syntactic priming effect in the control group.

A first possible explanation is the competitive nature of declarative and non-declarative/procedural memory systems (Krupa, 2009; Rieckmann and Bäckman, 2009). It has been found that these two memory systems are not strictly independent, but also interact with each other (Poldrack and Packard, 2003), and in the case of impairments in one system, the other might play a compensatory role (Ullman and Pullman, 2015). Indeed, animal studies have shown that the lesioning of one of the memory systems can result in an enhanced task performance relative to brain intact animals (Poldrack and Packard, 2003).

This results in the intriguing possibility that in the aging, healthy population the procedural memory contribution suffers from interference from the declarative memory system. Indeed, studies have suggested that certain aspects of priming, such as lexical overlap which we controlled for this in study, may be supported by explicit, declarative memory (Bock and Griffin, 2000; Chang et al., 2006; Ferreira and Bock, 2006; Bernolet et al., 2016) and hence recruitment of these systems provides an opportunity for competition. Damage to the declarative memory system would then result in removing the competition/interference, which in our case surfaces as a syntactic priming effect. Therefore, a combination of increased competition between memory systems for the healthy controls and an enhanced performance for the patients results in the large difference in priming magnitude that we observed in our study.

An alternative explanation is based on evidence from neuroimaging research, animal work, and patient studies that implicit memory depends on a subcortical-cortical network with particularly strong involvement of the striatum (Knopman and Nissen, 1994; Willingham and Preuss, 1995; Packard, 2009). As one ages, the putamen and caudate shrink by 5 – 10% (Raz et al., 2003) and dopamine in the striatum decreases as well (up to 10% per decade; Bäckman, Nyberg, Lindenberger, Li, and Farde, 2006). As the striatum is central in maintaining implicit information, the ability to maintain implicit information may also degrade. Currently only one study has looked at the effect of syntactic priming on aging and has suggested that, indeed, as we age our ability to prime decreases (Sung, 2015). Secondly, as we age the speed with which information is processed decreases (Howard, Heisey, and Shaw, 1986; Salthouse, 1996). Consequently, the chance that the information has decayed before it is retrieved is increased. In terms of syntactic priming, this could mean that the information is not retained long enough to be incorporated in future utterances. Indeed, one study has shown an increase in priming after administering a dopamine (via administration of levodopa) to healthy participants (Angwin et al., 2004).

The interesting question, however, is why do the amnesia patients in these studies not show a decrease in priming effect? One explanation may be that Korsakoff's patients have an increased 5-HT (a serotonin precursor) in the striatum (Langlais et al., 1987), which facilitates dopamine production (Zhou et al., 2005; Navailles and De Deurwaerdère, 2011). As the Angwin study mentioned above suggests, this increase in dopamine production may offer the Korsakoff patients better priming ability relative to their age-matched healthy peers.

In all, our results show unequivocally that syntactic priming is supported by implicit/procedural memory. Language processing, therefore, seems to rely not only on neocortically consolidated declarative memory, but also engages procedural memory structures, such as frontostriatal circuits, to engage in combinatorial processing, at least at the level of syntactic operations. To what degree reduced procedural memory contributions can be compensated by support from declarative memory remains to be seen.

3

In dialogue with an avatar, language behavior is identical compared to dialogue with a human partner

The use of virtual reality (VR) as a methodological tool is becoming increasingly popular in behavioral research as its flexibility allows for a wide range of applications. This new method has not been as widely accepted in the field of psycholinguistics, however, possibly due to the assumption that language processing during human-computer interactions does not accurately reflect human-human interactions. Yet at the same time there is a growing need to study human-human language interactions in a tightly controlled context, which has not been possible using existing methods. VR, however, offers experimental control over parameters that cannot be (as finely) controlled in the real world. As such, in this study we aim to show that human-computer language interaction is comparable to human-human language interaction in virtual reality. In the current study we compare participants' language behavior in a syntactic priming task with human versus computer partners: we used a human partner, a human-like avatar with human-like facial expressions and verbal behavior, and a computer-like avatar which had this humanness removed. As predicted, our study shows comparable priming effects between human and human-like avatar suggesting that participants attributed human-like agency to the human-like avatar. Indeed, when interacting with the computer-like avatar, the priming effect was significantly decreased. This suggests that when interacting with a human-like avatar, sentence processing is comparable to interacting with a human partner. Our study therefore shows that VR is a valid platform for conducting language research and studying dialogue interactions in an ecologically valid manner.

Adapted from: Heyselaar E, Hagoort P, Segaert K (2015) In dialogue with an avatar, language behavior is identical compared to dialogue with a human partner. *Behaviour Research Methods* 49(1), 46 - 60.

Introduction

The use of virtual reality (VR) as a method is becoming increasingly prevalent in behavioral studies in a wide range of fields, including navigation research (Tarr and Warren, 2002) and rehabilitation therapy (Rizzo and Kim, 2005). However this new trend does not seem to be catching on as strongly in the field of psycholinguistics. This may be due to the assumption that humans do not interact with computers in the same way that they interact with other humans, making any behavioral measure of language interaction with a computer partner ecologically equivocal. However, in this study we aim to debunk this assumption by showing that language processing in interaction with a human-like virtual agent ("avatar") is comparable to interactions with a human partner.

The assumption that humans do not interact with computers as if they have agency has already been shown to be false for interactions with desktop computers. Work by Nass and Moon (for review of their work see Nass and Moon, 2000) has repeatedly shown that humans attribute human-like characteristics to their desktop computer partner, the most unintuitive of these findings being the use of politeness when asked to evaluate the computer (Nass, Moon, and Carney, 1999). Participants would complete a task with Computer A and were afterwards asked to evaluate Computer A's performance. If Computer A conducted this evaluation, the ratings were significantly more positive than if the evaluation was conducted by another computer (or on paper), suggesting that participants were polite to Computer A. These behaviors were also replicated for other human-like traits such as the attribution of social hierarchy (Nass, Fogg, and Moon, 1996) and even ethnic stereotyping (Nass, Isbister, and Lee, 2000). These were all observed in participants who, during the debrief, agreed that "the computer is not a person and does not warrant human treatment or attribution." The authors suggest that this might be due to a phenomenon referred to as *Ethopoeia*: humans are inherently social and therefore human-like rules apply automatically and unconsciously also in interactions with computers. This phenomenon therefore would predict that language behavior should also be no different when conversing with a computer.

VR is one step up from desktop computers as it offers an almost real-world-like immersive experience that a screen and keyboard cannot offer. The reason we are focusing on VR is that it offers an immersive 3D world that participants can move in and interact with, allowing experimental control over parameters that cannot be (as finely) controlled in the real world, and only limitedly so in desktop computers. What is particularly important for interaction research is that VR offers the ability to finely control interlocutor behavior in parameters that are nearly impossible to control in a confederate, an aspect that is particularly attractive for language research.

Experimental studies of language behavior use conversation-like tasks that allow for manipulations of isolated specific conversation characteristics. These conversation-like tasks usually use a constrained conversation in which participants interact with either a naïve participant or a confederate. These experiments allow researchers to focus on, for example, the turn-taking event (Finlayson and Corley, 2012; Stivers et al., 2009), the role of the dialogue partner (Branigan, Pickering, Pearson, McLean, and Nass, 2003), and characteristics of the social interaction (Balcetis and Dale, 2012). In the turn-taking literature, more and more emphasis is being put on the role of subtle cues, such as intonational or lexico syntactic cues, to signal when a partner can begin preparing their response. Investigating the roles of these cues is complicated with a confederate, as the cues need to be exactly the same with each participant or manipulated to ensure accurate millisecond precision. Additionally, studies have shown that the opinion you have of your partner can influence how you comprehend and produce language. These social cues can be as subtle as interacting with an in-group member compared to an out-group member (Unger, 2010), having similar political views to your partner (Weatherholtz et al., 2014) or even participants that like the confederate more (compared to other participants) exhibit significantly different language behavior (Balcetis and Dale, 2012). Replacing confederates with a recorded message without any physical presence is therefore unnatural and the participants may not respond naturally to these cues. However, it is possible that if we introduce a human-like computer, one with rich human-like facial features and expressions, the language behavior of the participant might be natural enough to be comparable to language behavior in human-human interactions, yet allow fine enough control over the characteristics of the computer to allow for experiments that cannot be conducted with a human confederate.

In this study, we put VR as a methodology to study language behavior to the test. We focused on syntactic processing (specifying the syntactic relations between words in the sentence), a core aspect of language production and comprehension, in the form of the commonly used syntactic priming task. Linguistic priming refers to the phenomenon in which an individual adopts the language behavior of their conversational partner (e.g., different word choices, different syntactic structures, etc.; also referred to as *alignment* or *accommodation*, Pickering and Garrod, 2004). Syntactic priming, specifically, refers to adapting your sentence structure, or syntax, to match that of your partner and has also been indicated to be influenced by the opinions you have of your partner (see above). Therefore, a syntactic priming task is an ideal candidate to test whether VR is a valid replacement for human partners in conversation studies.

Replacing a human partner with a virtual agent or robot is not novel (Blascovich et al., 2002); many researchers investigate how participants interact with machines (Rehm and André, 2005; Bee et al., 2009; Pena et al., 2009; Rosenthal-von der Pütten et al., 2013; Melo et al., 2014). However, these studies only compare behavior towards different types of avatars but make no connection to “natural” behavior (i.e. comparing to participant behavior when interacting with a human in the same situations). Recently, there have been a few studies comparing human and virtual agent behavior in the language domain (Branigan et al., 2003; Pearson et al., 2006; Koulouri et al., 2014) which have shown that participants prime *less* with a human-like computer compared to a human partner. These studies use *belief* to convince participants that their partner is human/not-human. Unfortunately, a follow-up study has shown that language behavior in a belief condition does not match that of face-to-face language behavior (Bergmann, Branigan, and Kopp, 2015). The study had participants interact with a desktop computer. In certain conditions the participants were told that they were interacting with another human seated in another room, in other conditions they were told they were interacting with a program. The results showed a different language behaviour when the participants believed they were interacting with a human compared to a computer (similar to results shown by other human-computer language studies). However, if the participant was, in addition to the computer, presented with an animation of their apparent computer partner, their behaviour did match that of human-human interaction. The authors explain this as: “when social cues and presence as created by a virtual human come into play, automatic social reactions appear to override the intial beliefs in shaping lexical alignment” (pg. 9). Therefore, previous studies comparing human and computer interaction using only a belief manipulation may not be accurately measuring human-computer interactions. This emphasizes the importance of having an interlocutor present, which can be done using a desktop computer, but even more realistically when using VR.

In Experiment 1, we measured the magnitude of the priming effect when participants interacted with a human confederate and a human-like avatar with the hypothesis that priming behavior should be comparable. The human-like avatar had rich human-like facial expressions and verbal behavior. We conducted a follow-up experiment (Experiment 2) where participants interacted with the human-like as well as a computer-like avatar. The computer-like avatar had no facial expressions, her mouth movements did not match her speech, and all prosody was removed from her speech. As the magnitude of the priming effect is very susceptible to individual differences (Hartsuiker and Kolk, 1998b), both experiments were within-subjects designs so that we can measure how priming behavior *changes* as a participant interacts with different partner types.

We expect that the priming effect when interacting with the computer-like avatar will be significantly less, since we hypothesize that any comparable effects observed between the human and human-like avatar partner seen in Experiment 1 are due to the humanness of our human-like avatar. To prevent any influence of belief on the results, we did not tell the participants that they would be interacting with a human-like and computer-like avatar. Instead we only informed them that the avatars were speech-recognition programs and that the participants are participating in a pilot study.

As the previous literature has suggested that the magnitude of the priming effect can be influenced by the opinion one has of their conversation partner, we also investigated whether we could replicate those results here and whether the degree of this influence is comparable between human and avatar partners. Previous human-computer priming studies have not included this potential co-variate. Instead most have looked at a correlation between perceived conversational success and priming magnitude (unsurprisingly, as conversational success is measured as understanding each other (Pickering and Garrod, 2004), there is a positive correlation; Koulouri, Lauria, and Macredie, 2015). Therefore it will be interesting to see whether social influences on language behavior are also similar between human and avatar.

If we can provide evidence to support 1) that the magnitude of the priming effect is comparable between human-like avatars and humans, and that 2) these effects disappear when interacting with a computer-like avatar, we can confirm that VR is a valid method for replacing dialogue partners in language research. This could provide possibilities for new experiments in a wide range of subfields, such as turn-taking and social studies.

Experiment 1

In this experiment we investigate whether participants prime to the same extent with a human as with an avatar partner, and whether the magnitude of this priming effect is modulated by social factors as self-reported by the participants. This is to confirm that language interactions with a virtual partner are ecologically valid and that VR experiments can be used to replace future experiments with human partners.

Materials and Methods

Participants

53 native Dutch speakers gave written informed consent prior to the experiment and were monetarily compensated for their participation. Five subjects were not convinced that the confederate was an ignorant participant and/or did not believe that the avatar was voice-recognition controlled (see *Procedure*) and were therefore *a priori* not considered part of the data set. Thus only 48 were included in the analysis (21 male, M_{age}: 20.9; SD_{age}: 2.5).

Statistical Power

Statistical power was calculated using simulated priming data produced by the sim.glmm package (Johnson, Barry, Ferguson, and Müller, 2014) in R (R Core Development Team, 2011). For our simulated data set we assumed 25 repetitions per condition and 48 subjects. We assumed a 10% increase in passive production following a passive prime compared to baseline condition. With a difference between avatar and human priming magnitude of 6%, our simulated data set calculated a power of 0.751 with a 95% confidence interval of 0.722 – 0.777 after 1000 iterations.

Materials

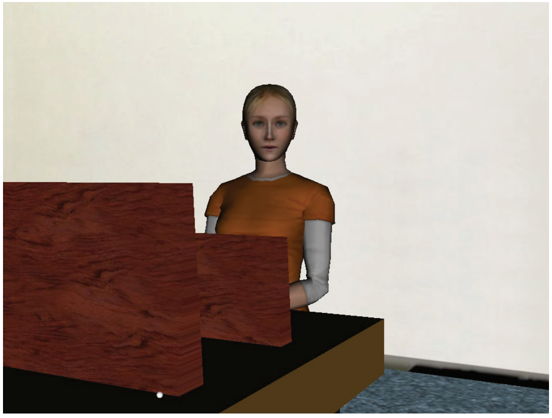

Figure 3.1 Avatar. The avatar exterior was identical for both avatar partners.

Avatar The avatar was adapted from a stock avatar produced by WorldViz ("casual15_f_highpoly"). All the avatar's speech was pre-recorded by a human female and played during appropriate sections of the experiment. The avatar's appearance suggested that she was a Caucasian female in her mid-twenties, which matched the age and ethnicity of the Dutch speaker who recorded her speech. This Dutch speaker was not the same as the confederate.

Avatar characteristics To choose the best and most human-like avatar, we collected data in a separate experiment (see Chapter 4). Six facial expressions (see Table 3.1) were judged by 30 participants not involved in the current study (13 male, M_{age}: 22.5; SD_{age}: 3.1) in categories such as humanness and strangeness, to see where they fell in the Uncanny Valley (Mori, 1970; Branigan, Pickering, Pearson, McLean, & Nass, 2003). The Uncanny Valley refers to the phenomenon in which human-like machines are perceived as more strange than their less human-like counterparts. While we wanted to select the most human-like avatar, we needed to ensure that said avatar does not cross this threshold.

Table 3.1 Characteristics of the 6 pre-tested avatars.

Avatar	Blink Duration[1]	Smiling Habit	Eyebrow Habit
1	No blink	No smile	No movement
2	0.5 seconds	1/(3-5 seconds)	No movement
3	0.5 seconds	Constant smile	Constantly up
4	0.1 seconds (Normal)	No smile	1/(3-5 seconds)
5	0.1 seconds (Normal)	Dialogue-matched	1/(3-5 seconds)
6	0.1 seconds (Normal)	Dialogue-matched	Dialogue-matched

[1]Measured from the beginning of the closing movement to when the eye is fully open again

Avatar 6 was rated as significantly more human and less strange than the other five (Tukey's HSD, $p < .05$) and was therefore used in our study (henceforth "human-like avatar"). This avatar would blink once every 1 – 5 seconds (blink duration was 0.1 second), raise her eyebrows once every 1 – 5 seconds, and when not speaking she would smile once every 5 – 10 seconds. During speech, her eyebrow and smile behavior was explicitly programmed to match the content of her speech. For example, the avatar would raise her eyebrows when asking a question and smile when she was enthusiastic (*Come, let's play another round!*). The influence of this blink and eyebrow behavior on agency is consistent with those of other studies looking into which characteristics support agency in the face (Looser and Wheatley, 2010; Tinwell, Grimshaw, and Williams, 2010), suggesting that these features (eyebrow movement and smile behaviour) are necessary features to create human-like avatars. Additionally, her mouth movement was controlled by a program that matched her mouth's openness to the pitch of the sound file. This created the illusion that her mouth movements were lip-synced to her speech.

Virtual environment The virtual environment (VE) was a stock environment produced by WorldViz ("room.wrl") adapted to include a table with a wooden divider (Figure 3.2). This divider was comparable to the physical divider used in the Human block. To ensure that the amount of time spent looking at the partner's face was the same between the Human and VE block, the divider was positioned so that while looking at the cards, the participant could not see the avatar's face unless they explicitly lifted their head. This is the same in the Human block. The reason the cards were not placed on the table itself in the VE is due to the weight of the head-mounted display (HMD), which would cause an uncomfortable strain on the back of the participant's heads when the participant faces down. Having the participants face forward distributes this weight more comfortably.

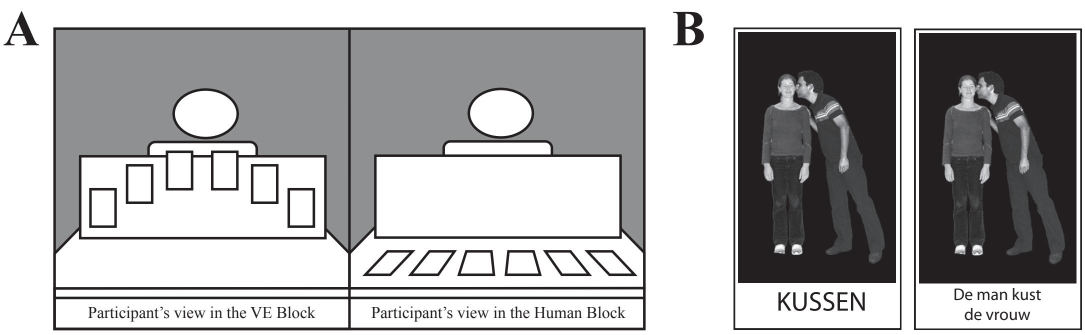

Figure 3.2 Set-up. A shows the experimental set-up from the view of the participant. The only difference is that in the VE the cards were presented at the top of the divider, whereas in the Human block, the cards were laid out on the table. **B** shows the participant card (left) and the confederate card (right). The participant card only showed the neutral verb associated with the photo, whereas the confederate card had a complete sentence written underneath. Here "to kiss" and "the man kisses the woman."

The table in the VE matched in both dimension and position with a table in the physical world, such that participants could physically touch the "virtual" table.

The experiment was programmed and run using WorldViz's Vizard software. Participants wore an NVIS nVisor SX60 HMD, which presented the VE at 1280x1024 resolution with a 60-degree monocular field of view. Mounted on the HMD were a set of 8 reflective markers linked to a passive infrared DTrack 2 motion tracking system from ART Tracking, the data from which was used to update the participant's viewpoint as she moved her head. It is known that this type of headset can cause dizziness and nausea due to the low frame-rate. However, for this study only 2 participants reported any discomfort. These participants were given a break and the option to continue; they both opted to complete the experiment. Additionally, a single reflective marker was taped onto the index finger of the participant's dominant hand. This marker was rendered as a white ball in the VE, such that participants knew the position of their finger at all times. Sounds in the VE, including the voice of the avatar, were rendered with a 24-channel WorldViz Ambisonic Auralizer System.

Stimulus pictures The pictures used in this task have been described elsewhere (Segaert, Menenti, Weber, and Hagoort, 2011). Our stimulus pictures depicted 40 transitive events such as *kissing, helping* or *strangling* with the agent and patient of this action. Each event was depicted by a greyscale photo containing either one pair of adults or one pair of children. There was one male and one female actor in each picture and each event was depicted with each of the two actors serving as the agent. The position of the agent (left or right) was randomized. These pictures were used to elicit transitive sentences; for each picture speakers can either produce an active transitive sentence (e.g. *the woman kisses the man*) or a passive transitive sentence (e.g. *the man is kissed by the woman*).

Filler pictures were used to elicit intransitive sentences. These fillers depicted events such as *running, singing,* or *bowing* using one actor. The actor could be any of the actors used in the transitive stimulus pictures.

Each card consisted of one stimulus picture with the relevant verb printed underneath. The cards were identical in the VE and Human block.

Questionnaire The questionnaire used in this study is adapted and translated from a questionnaire used in an earlier syntactic priming experiment by Weatherholtz and colleagues (2014). This study looked at the effect of political views on priming magnitude, and hence some questions were dropped as they were irrelevant for the current study ("My political views are usually conservative/liberal") or if they did not have a direct Dutch translation ("The speaker appeared intelligent" and "The speaker appeared smart" both translates to the same sentence in Dutch). The previous Weatherholtz study also looked at how participants deal with conflict situations, and whether that could have an effect on how much they adapt their own behavior to match that of their partner. We included that question set as well. Of these, all of the original English questions were included.

After each of the blocks (Human and VE), participants completed a questionnaire containing 7 6-point Likert-scale social evaluation items (hereafter *Relationship Questionnaire*, all questions listed in Table 3.2A). At the end of the experiment, participants filled in a questionnaire containing 7 6-point Likert-scale questions asking how they dealt with conflict (hereafter *Conflict Questionnaire*, all questions listed in Table 3.2B). All questions were phrased as statements, and the participants indicated the extent to which they agreed with each statement on a 6-point scale (6 = I absolutely agree, 1 = I do not agree at all).

Table 3.2A Factor loadings for the Relationship Questionnaire. Loadings greater than |0.4| are in bold as these items contribute most to the meaning of a factor. Loadings less than |0.1| are omitted for clarity.

	Factor 1	**Factor 2**	**Factor 3**
	Likeability	*Selflessness*	*Shyness*
I could be friends with my partner	**0.72**	0.37	0.19
My partner is similar to me	**0.73**	0.14	-0.13
My partner appeared generous	**0.53**	**0.62**	
My partner appeared intelligent	**0.84**	-0.12	**0.72**
My partner appeared selfish		**-0.92**	**0.87**
My partner appeared shy	0.15	0.21	**0.84**
My partner appeared enthusiastic	**0.53**	0.28	**0.72**
Proportion Explained	0.46	0.30	0.24

Table 3.2B Factor loadings for the Conflict Questionnaire. Loadings greater than |0.4| are in bold as these items contribute most to the meaning of a factor. Loadings less than |0.1| are omitted for clarity.

	Factor 4 *Ignore*	**Factor 5** *Dominance*	**Factor 6** *Compromise*
I ignored the conflict and behaved as if nothing had happened	**0.93**		
I pretended there was no conflict	**0.92**		
I tried to find a middle ground	0.13	-0.11	**0.89**
I had a discussion with the other person to try to find a middle groun	-0.30	0.25	**0.73**
I insisted that it wasn't my fault	0.13	**0.77**	
I kept pushing until the other person saw that I was right	-0.12	**0.83**	
I tried to convince the other person that my solution was the best	-0.23	**0.74**	0.20
Proportion Explained	0.37	0.37	0.27

Task and Design

All participants completed a language task probing syntactic processing in VE with an avatar (VE block) as well as in the physical world with a confederate (Human block; within-subjects design). The order of blocks was randomized and counterbalanced across participants. Partner type (human-like avatar vs. human) was used as an independent variable in the analysis.

Each block consisted of 228 trials (114 prime-target pairs). At the start of each block, the participant was presented with six cards, with the belief that the confederate/avatar had their own spread of six cards behind the divider (Figure 3.2). The participant and the confederate/avatar would alternate in describing cards to each other. If the listener saw the card that was described by their partner as one of the cards on the divider, then both conversation partners would remove that card from the spread and replace it with a novel card from their deck (in VE this would happen automatically after the card was identified). This continued until all 228 cards were described. The confederate/avatar description would *always* serve as the prime for the participants' subsequent target description.

The confederate's deck was ordered identically to the participant's deck, so the confederate/participant always had the card described to them. In the VE block, the avatar was programmed to randomly pick one of the participant's cards to describe thereby assuring that the participant always had the card described to them. If the participant described the card correctly (see *Procedure* below) the avatar/confederate admitted to having the card the participant described. The same cards were used for both partner types.

The confederate's deck of cards showed the stimulus picture but with a full sentence typed underneath, as such the confederate simply needed to read the sentence. 50% of the transitive sentences described the picture in the passive tense, 50% described it in the active tense. In VE, the avatar was programmed to use 50% passives, 50% actives.

The priming conditions were included in the analysis as independent variables. There were three priming conditions: baseline trials (intransitive prime followed by a transitive target), active priming trials (active prime followed by a transitive target), and passive priming trials (passive prime followed by a transitive target). However, as the participant was free to choose a card to describe, the chance existed that the participant would pick an intransitive card to describe in the target phase. These trials cannot be analyzed in terms of active or passive syntactic structure. Therefore, to ensure an adequate number of trials in each condition, out of the 228 cards 2/3 of the cards were transitive and 1/3 were intransitive. *Post-hoc* analysis showed that there was an average of 24.7 (SD: 7.4), 28.3 (SD: 3.4), and 25.3 (SD: 3.2) trials in the baseline, passive and active conditions respectively in the Human block and 20.7 (SD: 4.0), 24.6 (SD: 3.5), and 25.1 (SD: 4.0) trials in the baseline, passive and active conditions respectively in the VE block. One subject was discarded as the difference in the proportion of passive prime exposure between the two blocks (Human: 0.40; VE: 0.65) fell 2.5 standard deviations outside the mean difference between blocks (mean: 0.03; SD: 0.09).

Experimental Procedure

Participants were informed that our goal was to compare how experiencing events differed in VE compared to the real world. To ensure that the participants felt that they communicated with a program and not a programmer, they were told that it worked on voice-recognition, and hence no third party was necessary to operate the program. Questionnaires were handed out after each condition, as well as at the end of the experiment. Debrief questions were also handed out at the end of the experiment to see whether the participant believed the avatar to be independently operated and the confederate to be a naïve participant.

Responses during the syntactic priming task were manually coded by the experimenter as active or passive. An active sentence is one where the agent of the action is named first (e.g. *the woman kisses the man*) and such sentences were coded as 0; a passive sentence is one where the agent of the action is named last (e.g. *the man is kissed by the woman)* and were coded as 1. This way the data are ready to be entered into a logit analysis. An independent rater blind to the purpose of the experiment verified that the coding of a random sample of participants was done correctly (inter-rater reliability of 1). Target responses were included in the analysis only if 1) both actors and the verb were named correctly (as a sentence naming only one of the actors does not qualify as a transitive sentence) and 2) no unnecessary information was included in the description (which constrains the participants to using either an active or passive description). We excluded 1.0% (109 out of 10929) of the target responses because they were incorrect.

Coding and Analysis

Questionnaire As each participant filled in the Relationship Questionnaire twice (once for each partner type), we conducted a multivariate exploratory factor analysis on these results. For the Conflict Questionnaire, we conducted a principal components factor analysis. However, as this study only consists of a maximum of 199 data entries for the Relationship Questionnaire and only 95 for the Conflict Questionnaire, a value that is too low to conduct an accurate analysis, we combined our data with that of a similar study (Schoot, Hagoort, and Segaert, 2014)(Mori, 1970) to boost the total data set to 310 for the Relationship Questionnaire (Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO): 0.80; Bartlett's test of sphericity: $\chi^2(21) = 611.48$, $p < .0001$)), and 155 for the Conflict Questionnaire (KMO: 0.60; Bartlett's test of sphericity: $\chi^2(21) = 224.32$, $p < .0001$). This was possible as the study by Schoot and colleagues (2014) also investigated how priming magnitude changed across one experimental session, and used the exact same questionnaires as those used in this study. We used Jollife's criterion (Jollife, 1972, 1986) as the cut-off criterion (eigenvalues < 0.7) and extracted 3 factors per questionnaire. Table 3.2 shows the loading values for each of the extracted factors. Below each factor is the name we assigned to it, which we believe captures the theme of the factor best (i.e. the type of questions that contribute most to the meaning of the factor).

Mixed model analysis The responses were analyzed using a mixed-effects logit model, using the glmer function of the lme4 package (version 1.1.-4; Bates et al., 2012) in R (R Core Development Team, 2011). Target responses were coded as 0 for actives and 1 for passives. We used a maximal random-effects structure (Barr, Levy, Scheepers, and Tily, 2013; Jaeger, 2009): the repeated-measures nature of the data was modelled by including a per-participant and per-item random adjustment to the fixed intercept ("random intercept"). We attempted to include as many per-participant and per-item random adjustments to the fixed effects ("random slopes") as was supported by the data. The full model included random slopes for *Prime* for the per-participant and the per-item random intercept. The correlations between intercept and slope for these random effects were between -1 and 1, suggesting that the model has not been overparameterized. Likelihood ratio tests were used to ensure that the inclusion of the remaining random slopes and random intercepts are justified.

Partner Type was sum contrast coded (all means compared to the grand mean), *Prime* was dummy coded (all means compared to a reference group), and all numeric predictors were centered. We used the baseline (intransitive prime) as the reference group for *Prime*.

Results

Priming Magnitude is the Same Between Human and Human-like Avatar Partner

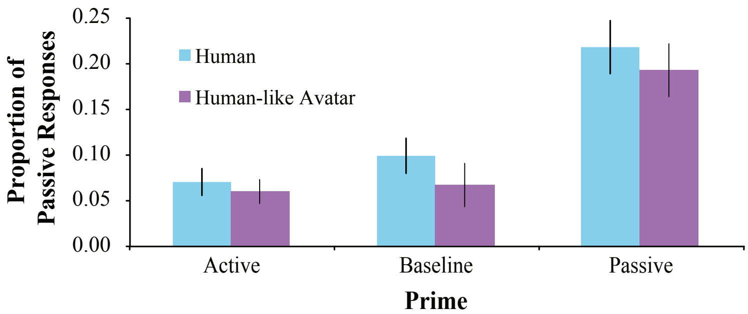

Figure 3.3 Proportion of passive responses per prime type for Experiment 1. As predicted, there are no significant differences in syntactic priming effects between the human and avatar block. Passive production increases with 11.8% for the human block and 12.3% for the avatar block following a passive prime compared to the baseline condition. In line with previous research, there were no priming effects for actives. Error bars represent standard error.

Figure 3.3 summarizes the relative proportion of passive target responses after each prime structure. To test our first hypothesis that the priming magnitude should not differ between partner types, we ran a basic logit mixed model with only *Prime* * *Partner Type* as a fixed effect. The output is shown in Table 3.3.

The negative estimate for the intercept indicates that in the baseline condition active responses were more frequent than passive responses. Following passive primes, more passive responses were produced compared to baseline (p < .001). Following active primes, there was no increase in active responses compared to baseline (p = .161). Neither active nor passive priming interacted with partner type (β = -0.12, p = .372; β = -0.11, p = .313 respectively) suggesting that the priming effect is the same in the Human and VE block. The main effect of *Partner Type* is marginally significant (p = .052).

Looking at Figure 3.3, this is most likely driven by the fact that there are marginally less passives produced when interacting with the avatar overall, regardless of prime type. Importantly, the priming effect is not significantly different between partner types (11.8% for human partner, 12.3% for avatar partner, p > .313).

Table 3.3 Summary for fixed effects in the mixed logit model for passive vs. active response choices between Human and Human-Avatar.

Predictor	coefficient	*SE*	*Wald Z*	*p*	
Intercept (intransitive prime)	-3.29	0.28	-11.89	< .001	***
Active Prime (AP)	-0.32	0.23	-1.40	.161	
Passive Prime (PP)	1.38	0.18	7.55	< .001	***
Partner Type (Human vs Avatar)	0.18	0.09	1.95	.052	
AP * Partner Type	-0.12	0.13	-0.89	.372	
PP * Partner Type	-0.11	0.11	-1.01	.313	
N = 6931, log-likelihood = -2004.6		* < .05	** < .01	*** < .001	

Influences on the Magnitude of the Priming Effect

To test our other hypotheses, we ran a mixed model in which we included all other measured variables, such as *Cumulative Passive Proportion, Order, Gender*, and all factors extracted from the questionnaire as well as interactions of all these factors with *Cumulative Passive Proportion*. *Cumulative Passive Proportion* is another way to present *Prime* but including a temporal element: the factor is calculated as the proportion of passives out of the total target responses produced before the current target trial. A positive *Cumulative Passive Proportion* therefore suggests that the proportion of passives previously produced positively influences the probability of producing a passive on the current target trial. *Order* and *Partner Type* were sum contrast coded, *Prime* was treatment coded (baseline primes as reference level), and numerical variables were centered. We started with a full model (AIC: 3979.2, BIC: 4232.5), and performed a step-wise "best-path" reduction procedure (using the drop1 function in R) to locate the simplest model that did not differ significantly from the full model in terms of variance explained (AIC: 3967.8, BIC: 4152.6, $p = .574$). The collinearity was low (VIF < 1.90). This best model is illustrated in Table 3.4.

Table 3.4 Summary of fixed effects in the best model of influences on passive priming between Human and Avatar partners.

Predictor	coefficient	SE	*Wald Z*	*p*	
Intercept (intransitive prime)	-3.76	0.25	-15.28	< .001	***
Order	1.06	0.28	3.82	< .001	***
Cum. Passive Proportion	3.33	0.43	7.82	< .001	***
Partner Type (Human vs. Avatar)	-0.00	0.06	-0.01	.993	
Likeability	0.27	0.09	3.05	.002	**
Dominance in Conflict	0.04	0.17	0.26	.794	
Dominance in Conflict * AP	-0.27	0.13	-2.04	.041	*
Dominance in Conflict * PP	-0.21	0.13	-1.62	.105	
Cum. Passive Proportion * Partner Type	-0.58	0.28	-2.07	.039	*
N = 6931, log-likelihood = -1956.9		* < .05	** < .01	*** < .001	

The model shows significant contributions to passive production from *Order, Cumulative Passive Proportion, Likeability,* and *Dominance*. We will address each of these contributions in turn.

Order The model shows a significant main effect of order. Specifically, there were significantly more passives produced in the second block (16.6% of all responses) compared to the first block (7.5% of all responses). This could be due to the fact that the participants have not interacted with an avatar before, and have also never completed the 6-card priming task before and therefore might require some time to get their bearings. As the partner types were counterbalanced, this does not influence our main findings.

Cumulative passive proportion The current priming literature suggests that priming occurs due to implicit learning as the proportion of passives produced by the participant increases as a function of time. *Cumulative Passive Proportion* was calculated as the proportion of passives out of the total transitive responses produced before the current trial; Figure 3.4 illustrates a similar increase over time for the human and avatar partner. Although the model shows that this effect is significantly different between partner types (p = .039) the trend of this effect is the same between partner type. The significant difference is therefore most likely driven by the fact that there are less passives produced overall with the avatar partner compared to the human partner.

We evaluated the influence of the *Cumulative Passive Proportion* on the syntactic response choice during the subsequent trial. Our mixed model analysis shows that a higher *Cumulative Passive Proportion* significantly increases the probability of a passive being produced. In other words, there is a cumulative

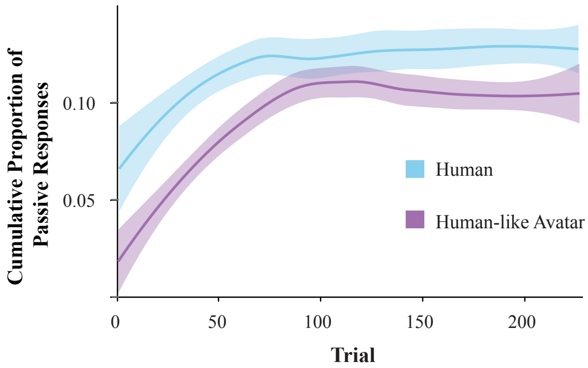

Figure 3.4 Cumulativity of passive responses for Experiment 1. The proportion of passives produced increases for all partner types over the course of the block. Mixed models show that there is a significant difference between the probability of producing a passive between human and avatar blocks (*p* = .445). This is most likely due to the lower starting point of the avatar partner. The learning curve (between trial 0 and 65) is equally steep for other partner types. there were no priming effects for actives. Error bars represent standard error.

effect of syntactic priming (i.e. the more passives produced, the stronger the effect) providing evidence for implicit learning as a mechanism for syntactic priming.

Likeability The questionnaire we used is directly adapted from the questionnaire used in an earlier syntactic priming experiment by Weatherholtz and colleagues (2014) and would thus be expected to provide the closest means of comparison. However, although Weatherholtz and colleagues (2014) observed a positive influence of *Likability* on priming magnitude, we could find no such influence: we only found a positive effect of *Likeability* on passive production, regardless of the prime type, such that the more the participant liked their partner, the more passives they produced. This effect was not significantly different for human vs. avatar.

It is surprising that the effect of *Likeability* on priming magnitude is qualitatively different between the two studies, even though comparable means of measuring liking was used. The effect of *Likeability* on priming magnitude should therefore be interpreted with caution. There are several differences in the set-up of the two studies, which suggests that social factors might mediate priming differently depending on the contextual and social environment, participant goals, etc. Further investigations will be necessary to elucidate how social factors influence priming magnitude.

Dominance in Conflict Our results also revealed a significant influence of *Dominance* on passive production: participants who rated themselves as being more dominating when dealing with a conflict produced less passives. Just like *Order*, the lack of a significant interaction with *Partner Type* indicates that this was independent of whether the participant's partner was human or computer. The model suggests that *Dominance* has a significant influence on the passives produced following an active prime, but not a passive. This is depicted in Figure 3.5. This figure shows a negative trend for passive production following both passive and active prime with increasing self-ratings of dominance in a conflict. Although the trend seems the same for both prime types, only active primes came out as significant in the model, most likely because the variability is lower

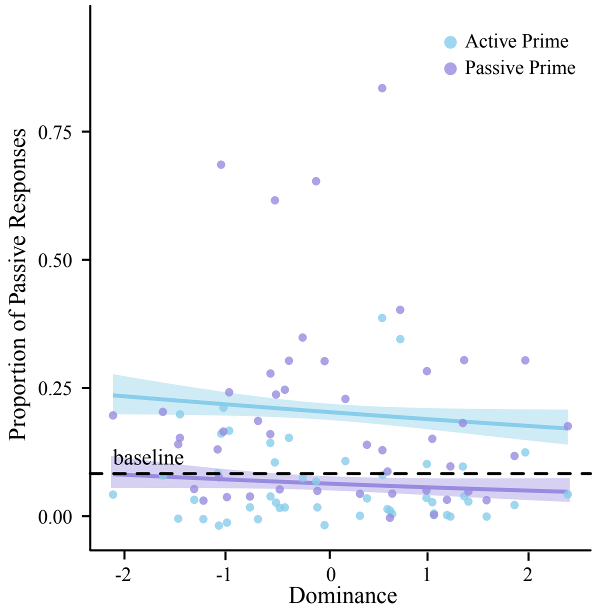

Figure 3.5 Effect of Dominance on passive priming. With increasing self-ratings of Dominance in Conflict, participants produced less passives compared to participant's who rated themselves as less dominant in conflict situations. The model stated that there is a significant difference between how responses are effected based on their prime type (active vs. passive; p = .041) however upon closer observation this effect may be influenced by the variability. Error clouds represent standard error.

(as indicated by the narrower error cloud).

Interim Discussion

Our data showed that our initial hypothesis was correct: priming magnitude is comparable between human and human-like avatar partners (12.3% vs. 11.8%). Investigating potential influences on priming magnitude revealed no significant effect of partner type suggesting that these trends are identical between human and human-like avatar partner. In terms of social influences, we did not see a difference between partner types, and most only influenced the overall proportion of passives produced, but not the priming effect itself.

We propose that this null-effect on priming magnitude is due to the humanness of our avatar, yet to be able to claim this unequivocally, we need to show that the priming effect is not present when the avatar used is not human-like.

Experiment 2

To determine whether the similarity in language behavior between the human and avatar condition is due to the perceived humanness of the avatar, we conducted a separate experiment in which language behavior was compared between a human-like and a computer-like avatar. In this experiment, the human-like avatar is the avatar used in Experiment 1. To create a computer-like avatar, we attempted to remove as much humanness as possible from the human-like avatar, i.e., it has no facial expressions, it doesn't look at the participant, and all prosody was removed from the audio files.

Materials and Methods

Participants

55 native Dutch speakers gave written informed consent prior to the experiment and were monetarily compensated for their participation. None of these participants took part in Experiment 1.

7 subjects did not believe that the avatar was voice-recognition controlled and were therefore *a priori* not considered part of the data set. Thus only 48 (22 male, M_{age}: 22.08; SD_{age}: 2.79) were included in the analysis.

Materials

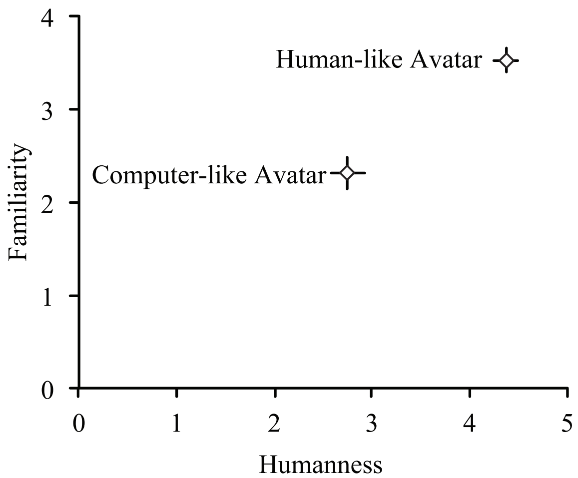

Figure 3.6 Humanness and Familiarity ratings of the two avatar types. Ratings were given immediately after the encounter with the avatar, although participants were able to change their answer after they had been exposed to both. Error bars represent standard error. The computer-like avatar was rated as significantly less human ($p < .001$).

Computer-like avatar The exterior of the avatar matched that of previous experiments. However, all facial expressions were removed and when she spoke her mouth no longer matched the pitch of her speech; instead it opened and closed in a loop, very much like a fish. The avatar was also programmed to stare straight ahead, instead of always looking at the participant. To ensure that participants do not respond to the humanness in the audio files, the pitch range was set to 0 in all audio files, which caused all prosody to be removed. Regardless, participants on average gave the computer-like avatar a rating of 4.5 (out of 6; SD: 1.25) on how easy she was to understand, compared to a 5.6 (SD: 0.54) for the human-like avatar.

To ensure that there is a difference between the human-like and computer-like avatars in terms of humanness, participants were asked to rate both avatars on their humanness and their familiarity. The results of this are shown in Figure 3.6, illustrating that there is a significant difference between avatars in the humanness category ($p < .001$, Paired *t*-test) but not in the familiarity category ($p = .220$, Paired *t*-test).

Task, Design, and Procedure

Task, design, and procedure matched the VE block of Experiment 1. Participants were also asked to rate the avatars using the questionnaire from Experiment 1. We again manipulated priming (baseline, active, and passive prime) and partner type (computer-like vs. human-like avatar) as independent variables. We measured and analyzed syntactic priming choices for the target sentences.

Post-hoc analysis showed that there was an average of 20.4 (SD: 5.5), 24.7 (SD: 3.5), and 25.0 (SD: 3.0) trials in the baseline, passive, and active conditions respectively in the human-like avatar block and 20.1 (SD: 6.5), 23.5 (SD: 3.5), and 25.2 (SD: 3.4) trials in the baseline, passive, and active conditions respectively in the computer-like avatar block. No participants needed to be excluded in this experiment due to unbalanced passive exposure between blocks.

Coding and Analysis

We excluded 0.65% (71 out of 10861) of the target responses because they were incorrect (criteria described under *Procedure* of Experiment 1). For the logit mixed model, the same procedures were used as in Experiment 1 except for this model we included *Prime* and *Partner Type* as random slopes for the per-item random intercept. The per-subject random intercept is the same as Experiment 1 (*Prime* as a random slope).

Results

Priming Effect Disappears with Computer-like Avatar

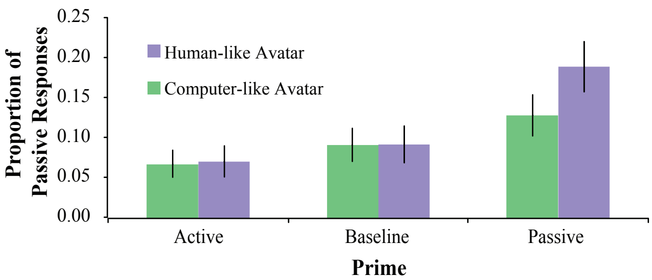

Figure 3.7 Proportion of passive responses per prime type for Experiment 2. There are significant differences in syntactic priming effects between the two avatar types (p = .033). Passive production increases by 9.5% with the human-like avatar and only 3.7% with the computer-like avatar following a passive prime compared to the baseline condition confirmed our prediction that participants primed less with the computer-like avatar as it is less human-like. In line with previous research, there were no priming effects for actives.

Figure 3.7 summarizes the relative proportion of passive target responses after each prime structure. The fixed effects of the model fit for these data are summarized in Table 3.5. To test our first hypothesis that the priming magnitude should be different between partner types, we ran a basic logit mixed model with only *Prime* * *Partner Type* as a fixed effect. The output is shown in Table 3.4.

Table 3.4 Summary for fixed effects in the mixed logit model for passive vs. active response choices between Human-Like and Computer-Like Avatar partners.

Predictor	coefficient	*SE*	*Wald Z*	*p*	
Intercept (intransitive prime)	-3.51	0.33	-10.49	< .001	***
Active Prime (AP)	-0.38	0.25	-1.50	.135	
Passive Prime (PP)	1.14	0.23	4.90	< .001	***
Partner Type	0.03	0.09	0.33	.740	
AP * Partner Type	0.00	0.13	0.01	.996	
PP * Partner Type	0.24	0.11	2.14	.032	*

N = 6627, log-likelihood = -1743.4 * < .05 ** < .01 *** < .001

The negative estimate for the intercept indicates that in the baseline condition active responses were more frequent than passive responses. Following passive primes, more passive responses were produced compared to baseline ($p < .001$). Following active primes, there was no increase in active responses compared to baseline ($p = .135$). As predicted, there was an interaction between passive priming and partner type ($\beta = 0.24$, $p = .032$) suggesting that participants primed less with the computer-like avatar compared to the human-like avatar.

Influences on the Magnitude of the Priming Effect

To test our other hypotheses, we ran a mixed model in which we included all other measured variables, such as *Cumulative Passive Proportion, Order, Gender*, and all factors extracted from the questionnaire as well as interactions of all these factors with *Cumulative Passive Proportion*. *Prime* was treatment coded with intransitive (baseline) primes as reference group. The remaining categorical factors were sum-contrast coded and all numerical variables were centered. We started with a full model (AIC: 3430.7, BIC: 3716.1), and performed a step-wise “best-path” reduction procedure (using the drop1 function in R) to locate the simplest model that did not differ significantly from the full model in terms of variance explained (AIC: 3416.0, BIC: 3572.3, $p = .226$). The collinearity was low (VIF < 1.54). This best model is illustrated in Table 3.5. The model shows significant contributions to passive production from *Cumulative Passive Proportion, Partner Type,* and *Dominance*. We will address each of these contributions in turn.

Table 3.5 Summary of fixed effects in the best model of influences on passive priming between Human-Like and Computer-Like Avatar partners.

Predictor	coefficient	*SE*	*Wald Z*	*p*	
Intercept (intransitive prime)	-3.54	0.27	-12.94	< .001	***
Active Prime (AP)	-0.29	0.22	-1.29	.197	
Passive Prime (PP)	1.19	0.22	5.28	< .001	***
Cum. Passive Proportion	4.87	0.52	9.44	< .001	***
Partner Type (Human-like vs. Computer-like)	0.19	0.06	3.28	.001	**
Dominance in Conflict	0.40	0.16	2.54	.011	*
Cum. Passive Proportion * Partner Type	-1.51	0.39	-3.19	< .001	***

N = 6607, log-likelihood = -1685.0 * < .05 ** < .01 *** < .001

Cumulative passive proportion Passive production over time is illustrated in Figure 3.8, again showing that the proportion of passives produced increases over the course of the block for both partner types. *Cumulative Passive Proportion* is a significant predictor of syntactic response choice, similar to Experiment 1. Also similar to Experiment 1 is the significant interaction of *Cumulative Passive Proportion* and *Partner Type* ($p < .001$). Again, however, the shape of the curves (i.e., the learning effect of priming) appears to be the same for the human- and computer-like avatars. Therefore, we believe this interaction is driven by there being less passives produced overall in the computer-like avatar condition.

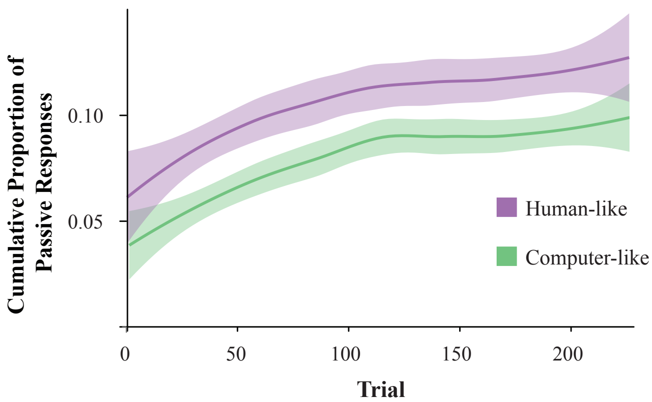

Figure 3.8 Cumulativity of passive responses for Experiment 2. The proportion of passives produced increases for both partner types over the course of the block. Mixed models show that there is a significant difference between the probability of producing a passive between the two avatar types ($p < .001$).

Partner type The main effect of *Partner Type* suggests that there are fewer passives produced when interacting with the computer-like avatar (9.7% of all responses) compared to the human-like avatar (11.8% of all responses).

Dominance in conflict Similar to Experiment 1, we find a main effect of self-rated *Dominance in Conflict* on passive production. This is illustrated in Figure 3.9A. Contrary to Experiment 1, in this experiment participants who rated themselves as more dominant in a conflict situation

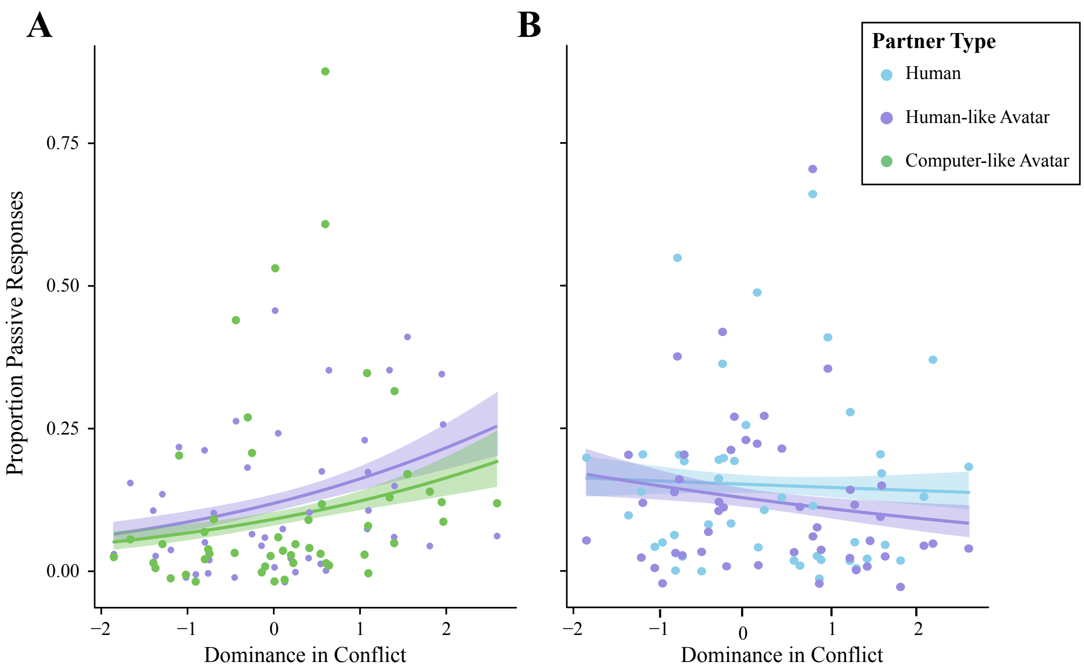

Figure 3.9 Effect of Dominance on passive production per partner type. A shows the effects for Experiment 2. As self-ratings of dominant behaviour in a conflict situation increase, the proportion of passives produced also increases. Curiously, **B** shows the opposite trend for Experiment 1. The human-like avatar is identical in both experiments, showing that this trend is most likely caused by the group make-up being different between experiments. This highlights the sensitivity of social factors to individual differences.

showed increased passive production compared to participants who rated themselves as less dominant. Although the model did not show a significant difference between partner types, to ensure that this flip in results is not due to the computer-like avatar, we plotted the results per partner type. This shows that this effect is equally strong for both human-like and computer-like avatar. Plotting the same for Experiment 1 (Figure 3.9B) again shows that it is not driven by one partner type; the results are exact opposites despite the conditions being identical (and the human-like avatar is identical) for the two experiments. This highlights individual differences in social factors and their influence on syntactic choice.

General Discussion

To validate whether VR is an ecologically valid method to study language in an interactive dialogue context, we measured syntactic processing during interactions with a human and two different avatar partners. To measure syntactic processing we performed the commonly used syntactic priming task and compared priming magnitude between the three conversation partners. Fully in line with our predictions, the results show comparable syntactic priming when participants interacted with a human partner compared to an avatar partner with rich human-like facial expressions and verbal behaviour (“human-like avatar”). When

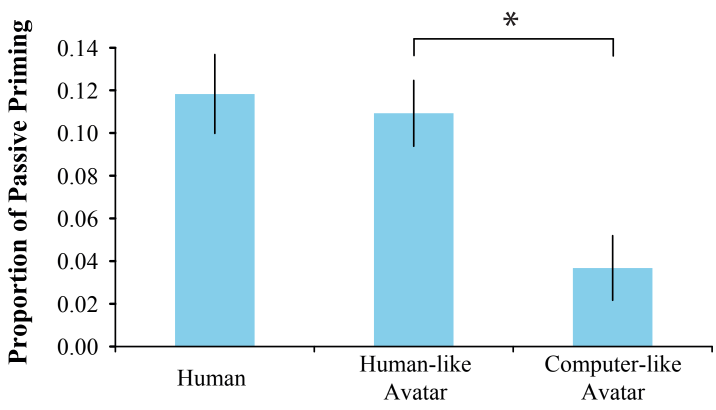

Figure 3.10 Priming magnitude per partner type. As priming with the human-like avatar was not significantly different between experiments (p = .850), the data was collapsed across experiments. Participants primed comparably with human and human-like avatar partners, but significantly less with the computer-like avatar (p =.030). As the only difference between the avatars was the humanness rating, the results suggest that the high priming magnitude seen with the human-like avatar is due to its perceived humanness.

participants interacted with an avatar partner with all this richness removed (no facial expressions and no prosody in speech; “computer-like avatar”) this comparable syntactic priming effect disappeared. Our results therefore suggest that participants who are interacting with a human-like avatar elicit the same language behaviour as if they were interacting with a human partner. We are attributing this finding to the humanness of the avatar, as when the experiment was repeated with an avatar that was rated as significantly less human compared to the human-like avatar, these effects were significantly reduced (Figure 3.10).

Three findings provide converging evidence that language behaviour was similar when interacting with the human-like avatar compared to the human partner: i) Syntactic priming effects were found when interacting with the human-like avatar as well as when interacting with the human partner and the size of these effects did not differ. In line with the literature, syntactic priming effects showed an inverse preference effect (syntactic priming effects for passives, not for active (Bock, 1986a; Ferreira, 2003)) and these again did not differ between the two partner types; ii) The influence of social factors on priming magnitude was not different between the human and human-like avatar partner; and iii) In line with the literature, the chance of producing a passive increased as a function of time, suggesting the presence of implicit learning in our task (Chang et al., 2006; Jaeger and Snider, 2008).

In this study we show that priming magnitude significantly deteriorates when interacting with a computer-like partner. However, we are not suggesting that a partner is necessary for the priming effect to take place. Indeed, previous studies have shown that priming occurs without the physical presence of a partner (Levelt and Kelter, 1982; Stoyanchev and Stent, 2009) and therefore a conclusion one could draw is that the lack of humanness of the computer-like avatar is acting as an interfering factor that prevents the participant from priming to the fullest extent. We suggest that if a study is in need of a physical presence, a human-like avatar can replace a human partner and will elicit human-like behaviour whereas a non-human-like avatar will most likely inhibit naturalistic behaviour. These things should be taken into consideration when designing the avatar partner.

In the current study, the human-like and computer-like avatar differ in that we used a computer-like or human-like voice, in addition to the use of facial expressions such as smiling and blinking habits. In this study we wanted to test two avatars that were as far removed as possible while still having them look identical. Therefore, based on these findings alone, we cannot conclude whether our findings are driven only by a difference in facial expressions, or also due to the use of different voices. However, in Chapter 4 we only manipulated the facial expressions (the voice was identical for all avatars) and we again find that perceived humanness determined syntactic priming effects: there was less priming for the avatar with no/less human-like facial expressions compared to the avatars with facial expressions.

In addition to looking at the differences in priming magnitude between the three partner types, we also investigated whether the same factors influence this increase in passive production behaviour. Although we show an influence of social perception and personality of the participant on influencing passive production, our results were not consistent between experiments. Although the set-up and methodology was exactly the same between experiments (including the human-like avatar) we found differences in which factors influenced and how they influenced passive production, namely the factors *Likeability* (how likeable the participant found their partner) and *Dominance in Conflict* (how dominant participants rate themselves when in a conflict situation). Although the influences of the factors differ between experiments, there was no significant influence of *Partner Type* on the magnitude or direction of the factors suggesting that this difference is purely due to the different participant groups used. This highlights the danger of using between-subjects designs to look at social influences, as the preferential make-up of one group does not always match the make-up of the other, as we highlight here. The big difference in the influence of social factors between experiments also highlights how susceptible these factors are to individual differences, even in groups of 48 participants.

Although this study only provides evidence for syntactic processing, it suggests the possibility that other language behaviours may also be consistent between VE and the real world. Syntactic processing is a core aspect of language, and occurs at a high level of sentence processing (Hagoort, 2005) suggesting that events that occur at earlier levels in language processing could also be tested using avatar partners. Indeed, evidence for speech rate and pitch adaptation with avatar partners has already been shown (Casasanto, Jasmin, and Casasanto, 2010). This opens pathways for the use of VR to investigate social behaviour in the field of psycholinguistics. With the commercialization of virtual reality machines, marketed for the average family (e.g. Oculus Rift) or anyone with a smart phone (e.g. the Cardboard app by Google), we believe that the current financial limitation of building a virtual reality lab will not be an issue in holding back future research. Additionally, studies have shown that as long as there is a virtual presence, even if that is an animated avatar presented on a desktop, the behaviour elicited by participants is comparable to their behaviour when interacting with another human, compared to a desktop without an animated being (Bergmann et al., 2015). VR offers a lot more possibilities for animation, but for some studies the animation possibilities of desktop computers would already be sufficient.

Additionally, our results also span into the field of robotics. Robotics has largely been concerned with creating human-like, realistic robots without investigating if humans interact with them the same as they would towards another fellow human (Segaert et al., 2011). Recent studies have already started to investigate which features of the robot are necessary to get users to attribute agency to them, and the results are consistent with what we have found in our current study: simple features are the key. For example, one study has shown that a robot will be rated as having agency because it cheats when playing simple games such as rock-paper-scissors or battleships (Short, Hart, Vu, and Scassellati, 2010; Ullman, Leite, Phillips, Kim-Cohen, and Scassellati, 2014). Our study can add to this new area that simple facial expressions such as random smile and eyebrow movement are enough to elicit human-like behaviour towards human-like robots. Future studies can use VR as an easily executable yet systematic method to determine which features are necessary to elicit agency.

In summary, VR provides an important platform on which previously unanswerable questions can now be investigated, providing a controlled method that produces results comparable to those seen in human literature.

4

How social opinion influences syntactic processing – an investigation using Virtual Reality

The extent to which you adapt your grammatical choices to match those of your interlocutor's (syntactic priming) can be influenced by the social opinion you have of your interlocutor. However, the direction and reliability of this effect is unclear as different studies have reported seemingly contradictory results. We have operationalized social perception as the ratings of strangeness for different avatars in a virtual reality study. The use of avatars ensured maximal control over the interlocutor's behaviour and a clear dimension along which to manipulate social perceptions toward this interlocutor. Our results show an inverted U-shaped curve in syntactic priming magnitude for passives as a function of strangeness: the participants showed the largest priming effects for the intermediately-strange, with a decrease when interacting with the least- or most-strange avatars. The relationship between social perception and priming magnitude may thus be non-linear. There seems to be a 'happy medium' in strangeness, evoking the largest priming effect.

Adapted from: Heyselaar E, Hagoort, P, Segaert K (2017) How social opinion influences syntactic processing – an investigation using Virtual Reality. *PLoS One* 12(4), e0174405

Introduction

Mimicry is one of those psychological behaviours that everyone has witnessed and/or produced themselves: at one point or another we have attempted to directly copy someone else's utterance (perhaps with a mocking tone), tried to mirror someone's movements, or attempted to put on an accent that wasn't our own. However, besides the conscious mimicry we engage in (or observe), there is also a wide range of mimicking behaviour we engage in without being aware of it. The most well-known example of this is of participants mimicking each other's foot-tapping behaviour (Chartrand and Bargh, 1999). This "chameleon effect" has been observed in a wide range of behaviours, from performance on an intelligence test to walking speed (Bargh et al., 1996; Dijksterhuis et al., 1998; Dijksterhuis and Bargh, 2001). What makes this behaviour so interesting is that these participants were not aware they were changing their behaviour, and were primed by subtle manipulations such as using key terms hidden in scrambled sentences or questionnaires. Studies investigating this effect have also shown that not everyone mimics to the same extent (Lott and Lott, 1961). This variability can be attributed to, among other things, characteristics of the social interaction, such as the likability of the person one interacts with. Indeed, there have been studies suggesting that participants are more likely to mimic people they like, compared to those they dislike (Charny, 1966).

Although there has been some controversy on whether these results are replicable (Bargh, 2012; Bowers, 2012), these studies have inspired others to investigate whether this mimicking behaviour extends to other domains. Indeed, this behavioural priming effect can also be observed in the language field: Participants adapt their speech rate and accent (Giles and Powesland, 1975; Giles et al., 1992), even lexical (Bock, 1986b), and grammatical (Bock, 1986a; Bock et al., 1992; Hartsuiker and Kolk, 1998b) preferences, to name but a few, to match their interlocutor. Studies have shown that social factors can also influence the rate of language mimicry (for review see Giles et al., 1991) however these studies have only looked at the influence of social factors on what can best be described as superficial language traits: body language/posture (Maurer and Tindall, 1983), speech rate (Street et al., 1983), vocal intensity (Natale, 1975), etc. For higher level language change, such as lexical choice and syntactic structure, which represent changes in actual language processing, it is established that there is a priming effect, but less research has investigated whether there is an influence of social factors.

As language is a social behaviour, the suggestion that the opinion one has of their interlocutor could influence one's language processing, also at the core levels of semantic and syntactic processing, is not surprising. Although not focusing on mimicry, there have been several empirical studies that have illustrated that language processing, at the level of semantics as well as grammar, can be affected by social information. For example, participants show different brain patterns when statements were incongruent with inferred speaker characteristics

(Berkum et al., 2008; Tesink et al., 2009), or slowed reaction times when the grammatical gender of a sentence does not match the gender of the speaker (Andonova, 2013; Vitevitch et al., 2013). These studies suggest that when processing language, we already take speaker characteristics into account, and hence the assumption that something as abstract as social opinion could also have an influence is not as far-fetched.

For the rest of this paper we will focus exclusively on the adaptation of syntactic structure choices to an interlocutor's, and refer to it as structural adaptation or structural priming.

In 2005, Balcetis and Dale set out to answer whether syntactic adaptation is influenced by social factors by manipulating the opinion participants had of their interlocutor. Participants and confederates were invited to complete a picture description task together, with the aim of measuring how often the participant would mimic the grammatical structures used by the confederate. Before the experimental session began, the participant and confederate first completed a questionnaire so that they could "get to know each other" (p. 185). The manipulation was such that the confederate would answer the questions in either a mean or nice way. The questionnaires were exchanged such that each could read the answers of the other and, at the request of the experimenter, the participant and confederate stated what they thought of the other. To ensure that the "mean" confederate was indeed regarded as such, when in the mean condition, the confederate would answer that they would not be friends with the participant, whereas the "nice" confederate (although both styles of confederate were played by the same person) would say that they could be friends and thought the participant was "ambitious and exciting" (p. 186). They would then conduct the picture description task together.

This study showed that participants were more likely to mimic the syntax of the sentences produced by the nice confederate compared to the mean confederate for 3 out of 4 structures measured. Prepositional-object (PO), active, and passive sentences all showed a significant increase in the magnitude of the priming effect with the nice confederate compared to the mean one. Only double-object (DO) structures showed a (non-significant, $p = .13$) opposite correlation. Interestingly, in Experiment 2 of the Balcetis and Dale study, the opinion of the confederate was no longer manipulated by how they profiled themselves, but rather by how they interacted with the experimenter. The mean confederate in this experiment would complain when the experimenter pretended to have difficulty in setting up the experiment, while the nice confederate would be patient and understanding. Even with this more subtle manipulation, an independent group of participants rated the confederate's behaviour as significantly different. In this study, the participant mimicked more sentences with the mean confederate compared to the nice confederate, even for DO structures. These results highlight not only how even small changes to the perception of the confederate can largely influence the results but also puts into question the stability of the results in general. Is there a consistent and stable influence of social opinion on language mimicry?

In 2014, Weatherholtz and colleagues conducted a similar experiment, but instead of the binomial “mean” versus “nice” confederate they had participants complete a survey which measured how similar the participants found themselves to their partner, giving a wider spread of ratings. The experiment was conducted online, with participants first hearing a political diatribe after which they were asked to describe 10 simple line drawings. The authors manipulated the political ideology of the diatribe (political ideology of the participant was also measured in the survey) as well as the accent of the speaker. In this study, Weatherholtz and colleagues showed a decrease in PO priming with increasing similarity scores, and a non-significant (p = .3) increase in DO priming with increasing similarity scores. Assuming that we like people more if we find ourselves similar to that person, we can consider the results of the Weatherholtz study to be contradictory to the results of Experiment 1 of Balcetis and Dale, yet in line line with Balcetis and Dale’s Experiment 2. This calls into question what Balcetis and Dale, and Weatherholtz and colleagues were actually measuring. The Weatherholtz study did not have their participants take turns with their confederate in describing pictures; instead they were given ten pictures to describe in a separate experimental phase. As their results are more in line with Balcetis and Dale’s Experiment 2, in which there was also no direct manipulation between the confederate and the participant, perhaps these results are more in line with mimicry in a monologue context (i.e. without the direct influence of the interaction partner).

Regardless, both studies clearly provide evidence that social factors can influence language mimicry at the level of syntactic processing or structure choices, however, the direction of this effect is not clear. Social factors are inherently noisy due to the abstract and personal nature of this type of measurement. However, we believe that the manner in which the previous studies represented the interlocutor (confederate or online recording) may have also added noise to the data, potentially resulting in the conflicting results.

In language priming studies, it is important that the participants are equally exposed to all grammatical structures that the experiment is attempting to prime, and therefore a scripted partner (whether it be a confederate or a recording) is necessary to ensure enough exposure to each eligible structure. However, the use of a confederate has recently been scrutinized (see Kuhlen and Brennan, 2013 for a review) in terms of the potential for artificially produced signals that may influence the behaviour exhibited by the participant. In most priming studies this is not necessarily an issue, as all that is necessary is to observe an effect of priming on the participant’s syntactic choices. However, when using a confederate to manipulate the social factors towards that confederate, even uncontrolled minute changes in the confederate’s behaviour could influence the participant’s behaviour, as was illustrated in the diverging results between Experiment 1 and Experiment 2 in the Balcetis and Dale study, for example.

Weatherholtz and colleagues addressed these issues by replacing the confederate with a recording. In addition to controlling for minute changes in behaviour between participants, this also changed the social dynamic as the participants were not directly interacting with the confederate, just listening to a recording. Even the picture description task was just the participant describing pictures; there was no retort from a confederate. This made the manipulation subtler but also potentially changed the way the manipulation affected language production, as this was not production within the social context created. Additionally, in terms of linguistic priming, a recent study by Bergmann and colleagues has shown that participants adapt their language behaviour significantly *less* when presented with only a voice recording of their partner, as opposed to a voice recording and accompanying video (Bergmann et al., 2015). This study illustrated that when participants were interacting with a computer without a video of their partner, the priming magnitude was significantly different depending on whether the participant believed their partner to be a computer program (more priming) or a human (less priming) in another room. However, once a video of their partner was added, linguistic priming magnitude (they tested both lexical and syntactic priming) was no different between belief conditions. This leads to our approach in addressing the question of the relationship between social factors and language mimicry: By use of a computer with video.

In Chapter 3, we showed that participants adapt their grammatical structures to the same extent when interacting with a program in virtual reality (hereafter "avatar") as when they are conducting the same task in the physical world with a confederate (Chapter 3 this thesis). In this study, participants completed a picture description task together with an avatar or a human confederate. The target pictures could be described using an active or a passive sentence, and our results showed an increase in the use of passive structures following a passive prime that was of the same magnitude regardless whether the partner was an avatar or a human confederate. Crucially, this priming magnitude became marginal when participants interacted with a computer-like avatar: This avatar did not have any facial expressions, did not look at the participant and had a computerized voice. We interpreted these results as indicating that if the computer partner is human-like enough, then participants will exhibit the same behaviour towards it as they would a human partner, a claim supported in the human-computer literature (Gong, 2008; Lee, 2010).

These results, together with those of Bergmann and colleagues mentioned above, suggest that avatars in virtual reality could be a viable replacement for confederates, particularly in investigating the role of social factors. Indeed, this idea is not new (Blascovich et al., 2002) and although the effect of social factors influencing language production has not been investigated using computers, a large repertoire of studies from the Clifford Nass lab in Stanford as shown that participants do attribute social factors to machines (see for review Nass and Moon, 2000), and that these participants have indicated that they believe that the computer will reciprocate these feelings. Indeed, they conducted several tests on reciprocity: if a computer is particularly helpful (in terms of search ability; Fogg and Nass, 1997) or discloses sensitive information ("I rarely get used to my full potential. What is your biggest disappointment in life?"; Moon, 2000), participants respond by being helpful towards the computer (in helping the computer with programming a task) or disclosing sensitive information, compared to switching to a different (yet identical) computer. These studies have suggested that participants exhibit social characteristics towards computers and therefore it is not unfair to predict that the same will occur for language mimicry.

Hence the aim of this paper is two-fold: 1) we aim to show that social factors can influence syntactic priming when interacting with avatars and 2) we aim to determine how social factors can influence syntactic priming, with the claim that this should reflect the directionality in human-human studies as well. In Experiment 1 we detail how we picked our three avatars. Previous studies have shown that participants prime less with computer partners (avatars included) when they are perceived as being less human (Beckner et al., 2015; Chapter 3 this thesis). We will therefore create and test six avatars and pick three that are rated as least human, most human, and intermediate. These three human-like avatars will then be used as partners in a within-subjects structural priming task (Experiment 2), looking at the rate of priming for active and passive structures as a function of social opinion. We base our manipulation on the methodology employed by Weatherholtz and colleagues: A subtle manipulation by changing the behavioural features of the avatar, not an explicit manipulation as was done by Balcetis and Dale. We believe that this will more realistically reflect how social factors could influence language production in everyday life. Therefore, we will use the same survey used by Weatherholtz and colleagues to determine interpersonal similarity to each avatar as the independent variable against which to correlate priming magnitude.

This study will therefore answer how social factors influence language adaptation. We predict, based on the scarce previous literature, that in a social context (as in Balcetis and Dale's Experiment 1) we should see an increase in syntactic priming with increased similarity/likeability of the avatar. If we can unequivocally show this effect, then language adaptation would reflect the patterns seen in social psychology for behaviour priming, suggesting a similar, non-language specific, mechanism is being used in processing syntax.

Experiment 1

In Experiment 1, we tested six avatars with the aim to pick three that represent the least-, most-, and intermediately-rated human-like avatar. Our six avatars had unique combinations of different facial expressions, namely the blink rate, eyebrow movement, and smile habits. We chose to manipulate facial expressions only as, due to the nature of the syntactic priming task, the avatar does not move or say anything other than describing pictures. As such, we were limited in the behavioural characteristics that we could manipulate.

The participants conducted a shortened version of the syntactic priming task to ensure that the ratings are more relatable to Experiment 2. The participants were asked to rate the avatars on their perceived humanness, strangeness, quality of facial expressions, and quality of voice. The latter will be used as a sanity check as the voice was the same for all 6 avatars. Humanness and strangeness were included to ensure that the combinations we chose are realistic human-like expressions and do not evoke an unsettling feeling that may bias the participants structural priming behaviour as measured in Experiment 2.

Materials and methods

Participants

30 native Dutch speakers (13 men, M_{age}: 22.5 years; SD_{age}: 3.1) gave written informed consent prior to the experiment and were monetarily compensated for their participation. The study was approved by the ethics commission of the Faculty of Social Sciences at Radboud University, Nijmegen (Ethics Approval # ECG2013-1308-120).

Materials

Avatars. All avatars had the same exterior adapted from a stock avatar produced by WorldViz (“casual15_f_highpoly”). All the avatars’ speech was pre-recorded by the same human female and played during appropriate sections of the experiment. The avatar’s appearance suggested that she was a Caucasian female in her mid-twenties, which matched the age and ethnicity of the Dutch speaker who recorded her speech.

The six facial expressions to be tested involved combinations of subtle changes in blink rate, smiling, and eyebrow habits (Table 3.1). Blinks happened once every 1 - 5 seconds. For versions with dialogue-matched smiling and dialogue-matched eyebrow habits we explicitly programmed when the avatar would smile and/or raise her eyebrows, such that it would coincide with the content of her speech. For example, the avatar would raise her eyebrows when asking a question and smile when she was enthusiastic. When not speaking, she would smile once every 5 - 10 seconds and raise her eyebrows once every 1 - 5 seconds such that she would still differ from the no smile/no eyebrow version. All of these changes were extremely subtle to ensure that they can still be related to ecologically valid behavioural characteristics that one would encounter in the everyday world.

Table 3.1 Avatar Facial Expressions

Avatar	Blink Duration[1]	Smiling Habit	Eyebrow Habit
1	No blink	No smile	No movement
2	0.5 seconds	1/(3-5 seconds)	No movement
3	0.5 seconds	Constant smile	Constantly up
4	0.1 seconds (Normal)	No smile	1/(3-5 seconds)
5	0.1 seconds (Normal)	Dialogue-matched	1/(3-5 seconds)
6	0.1 seconds (Normal)	Dialogue-matched	Dialogue-matched

[1]Measured from the beginning of the closing movement to when the eye is fully open again

Virtual environment The virtual environment (VE) was a stock environment produced by WorldViz ("room.wrl") adapted to include a table with a wooden divider. We chose to have the cards displayed at the top of the divider so that the participants could see the cards while facing forward. This was done due to the weight of the head-mounted display (HMD), which would cause an uncomfortable strain on the back of the participants' heads when they face down. Having the participants face forward throughout the entire experiment distributes this weight more comfortably.

The experiment was programmed and run using WorldViz's Vizard software. Participants wore an NVIS nVisor SX60 HMD, which presented the VE at 1280 x 1024 resolution with a 60-degree monocular field of view. Mounted on the HMD was a set of 8 reflective markers linked to a passive infrared DTrack 2 motion tracking system from ART Tracking, the data from which was used to update the participant's viewpoint as she moved her head. It is known that this type of headset can cause dizziness and nausea due to the exclusion of the participant's nose in the field of view (Whittinghill et al., 2015). However, as each experimental block was quite short (~5 minutes), none of our participants reported feeling any nausea.

Additionally, a single reflective marker was taped onto the index finger of the participant's dominant hand. This marker was rendered as a white ball in the VE, such that participants knew the position of their finger at all times. Sounds in the VE, including the voice of the avatars, were rendered with a 24-channel WorldViz Ambisonic Auralizer System.

Procedure and Task

The participants were informed that they would be rating six different avatars. Exposure to each avatar started with the avatar giving a short introductory speech, followed by a card matching game. The order of the avatars was randomized and counterbalanced across participants, such that each participant interacted with all six types of avatar in all possible order combinations.

The card game is identical to the one used in Experiment 2 (for more details see *Materials and methods* of Experiment 2). The participant and the avatar would alternate in describing picture cards to each other. If the listener saw the card described by their partner as one of the cards in their spread they would select it, causing it to be automatically replaced by a novel card. The listener would then become the speaker and pick a card to describe. This continued until 10 cards were described, after which the headset was removed and participants were asked to fill out a pen-and-paper questionnaire. We favoured a pen-and-paper questionnaire instead of having the avatar ask the questions directly as previous research has shown that if the participant evaluates the avatar in the presence of said avatar, they rate them more favourably (Nass et al., 1999). The task was too short, however, to measure stable priming tendencies.

The questionnaire consisted of four 6-point Likert-scale questions asking to rate the avatar on perceived humanness, strangeness, quality of their facial expressions, and quality of their voice in relation to the other avatars. The latter is a sanity check as the voice is the same for all six avatars. The scale was such that 1 referred to least human/least strange/lowest quality, whereas 6 referred to most human/most strange/highest quality After each avatar, the participants were allowed to change their ratings for previously viewed avatars.

Results and Discussion

We found a significant effect of avatar versions on the rating of humanness (F = 4.970, $p < .0001$), strangeness (F = 3.065, $p = .01$; Figure 4.1), and quality of facial expression (F = 5.097, $p < .001$). The voice ratings were not found to be significantly different between avatar versions (F = 1.418, $p = .220$), which functions as a sanity check as the voice was exactly the same for each avatar.

A *post hoc* Tukey HSD test revealed that avatars with eyebrow movement (Avatars 3 - 6) were rated significantly more human than avatars without

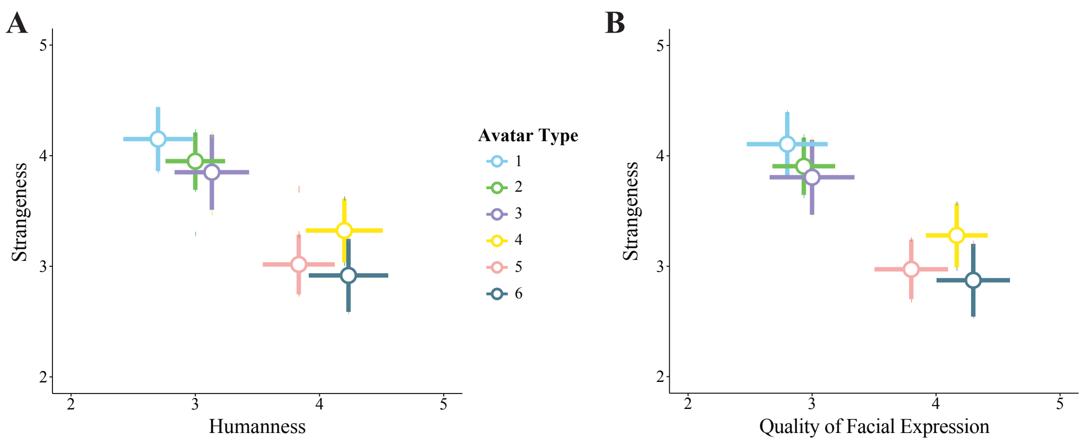

Figure 4.1 Rating of Avatar Versions for Experiment 1. A. Correlation between the strangeness and humanness ratings, and **B.** Correlation between the strangeness and quality of facial expressions for the six avatars. Error bars represent standard error.

eyebrow movement (Avatars 1 and 2, $p < .05$), whereas smiling habits made no significant impact on humanness rating (Avatars 1 and 4 vs. Avatars 2,3,5 and 6). This result is consistent with previous literature showing that the eyes are used to determine agency in humanoid objects (Looser and Wheatley, 2010). Those experiments used inanimate photographs as their main stimuli, and therefore it is interesting that we were able to replicate those effects here with animated, interactive beings.

The aim of this experiment was to pick three avatars that are the least human, most human, and an intermediate without being rated as overly strange. As all six were rated under 50% for strangeness, we chose the avatars with the lowest humanness rating (Avatar 1), highest humanness rating (Avatar 6), and one intermediate (Avatar 5).

Experiment 2

In this experiment, a new set of participants were invited to complete the full syntactic priming experiment with the three avatars that we chose based on the ratings obtained in Experiment 1. Each participant interacted with all three avatars. The within-subjects manipulation ensures that we avoid a possible confound of the individual differences between participants commonly seen in priming studies (Hartsuiker et al., 2008). For each avatar, participants were again asked to rate the avatar on humanness and strangeness, as well as complete a questionnaire evaluating the interpersonal distance with the avatar, based on the questionnaire used in the Weatherholtz and colleagues study. The aim of the current experiment is to determine whether these ratings correlate to the magnitude of the syntactic priming effect, to show how social perception can influence priming behaviour.

Materials and methods

Participants

66 native Dutch speakers (24 men; M_{age}: 21.08 years, SD_{age}: 2.179) who had not participated in Experiment 1 were invited to complete a syntactic priming experiment with each of the three avatars. All participants gave written informed consent and were monetarily compensated for their participation.

Statistical Power

Statistical power was calculated using simulated priming data produced by the sim.glmm package (Johnson et al., 2015) in R (R Core Development Team, 2011). For our simulated data set we assumed 15 repetitions per prime type (active, passive, baseline, see below). We assumed a 10% passive priming effect (10% more passives produced following a passive prime compared to an active prime), which is the order of magnitude commonly seen in the literature (Segaert et al., 2011; Chapter 3 this thesis). We simulated a maximum difference of 6% (based on the results seen in Chapter 3) between one end of the social perception scale and the other. With 66 participants, this would give our study a power of 0.957 (0.9414 0.9685; 95% confidence interval).

Materials

Avatar Avatar 1, 5, and 6 (see Experiment 1) were used to represent the least, intermediate, and most human avatar, respectively.

Virtual environment is the same as the one described in Experiment 1. Although the current experiment is longer than the one in Experiment 1, none of the participants reported any feeling of nausea.

Stimulus pictures The pictures used in this task have been used previously (Segaert et al., 2011). Our stimulus pictures depicted 40 transitive events such as *kissing*, *helping* or *strangling* with the agent and patient of this action. Each event was depicted by a grey scale photo containing either one pair of adults or one pair of children. There was one male and one female actor in each picture and each event was depicted with each of the two actors serving as the agent, creating four possible combinations for each event (160 transitive pictures in total). The position of the agent (left or right) was pseudorandomized such that across participants a certain card would be presented with the agent on the left and the right equally. These pictures were used to elicit transitive sentences; for each picture speakers can either produce an active transitive sentence (e.g. *the woman kisses the man*) or a passive transitive sentence (e.g. *the man is kissed by the woman*).

Filler pictures were used to elicit intransitive sentences. These fillers depicted events such as running, singing or bowing using one actor. The actor could be any of the actors used in the transitive stimulus pictures. There were 80 filler cards for the program to choose from.

Each card consisted of one stimulus picture with the relevant verb printed underneath.

Questionnaire After an interaction with each avatar, participants completed two questionnaires. The first is an Avatar Evaluation questionnaire identical to the one used in Experiment 1. For this questionnaire participants were asked to rate the avatars on a 6-point Likert scale on humanness, strangeness, quality of facial expression, and quality of voice in relation to the other avatars. The scale was such that 1 referred to least human/least strange/lowest quality, whereas 6 referred to most human/most strange/highest quality. The second was 7 questions relating to their social opinion of the avatar (adapted from Weatherholtz et al., 2014; hereafter *Interpersonal Distance Questionnaire*). These questions were phrased as statements (see Table 4.2 for a complete list) and the participants indicated the extent to which they agreed with each statement on a 6-point Likert scale (6 = *I absolutely agree*, 1 = *I do not agree at all*).

Task and Design

All participants completed a syntactic priming task in VR with each avatar. The experiment was split into three blocks: each block included the syntactic priming task with an avatar plus the Avatar Evaluation and Interpersonal Distance Questionnaire which were given after the priming task was complete. After the participant had completed all blocks, they were presented a debrief form (see below). The order of the blocks was randomized and counterbalanced across participants, such that each participant interacted with all three types of avatar in all possible order combinations.

The task is adapted from the syntactic priming task done in VR in Chapter 3. The participants were instructed to describe cards alternately with the avatar. Each block consisted of 150 trials (75 prime-target pairs) randomly picked from the database of 240 cards. At the start of each block, the participant was presented with six cards, with the belief that the avatar had her own spread of six cards behind the divider. The participants were instructed to choose one card per turn to describe and to describe it using a single concise sentence (e.g., *the man kisses the woman*). If either conversation partner had the card that was described, the card was automatically removed from the deck after the listener selected it. In truth, the avatar had an identical deck to the participant and therefore always had the card, however, would only indicate so to the participant if the card description met the necessary criteria (see *Analysis*). The six-card design was used as it creates the illusion that the avatar can understand the participant (they were told that the avatar works with a speech-detection system). This will ensure that the participant is priming with each avatar as an individual and not with the program (for a discussion on this see Nass and Moon, 2000). All participants were asked during the debrief whether they believed in this manipulation; if not, they were not included in the data set (5 out of 71 participants were discarded *a priori* for this reason).

The avatar was programmed to randomly pick one of the participant's cards to describe, thereby assuring that the participant always had the card described to them. The avatar was programmed to use 50% passive descriptions, 50% active descriptions. At the beginning of the block, the avatar would always go first, thereby serving as the prime for the participants' subsequent target descriptions.

There were two priming conditions: active priming trials (active prime followed by a transitive target) and passive priming trials (passive prime followed by a transitive target). There was also a baseline condition: the avatar would describe an intransitive card (thereby no using an active nor passive structure) and the participant would respond with a transitive card. This condition was used to measure the tendency for the participant to use active and passive structures without being primed.

However, as the participant was free to choose a card to describe, the chance existed that the participant would describe an intransitive card as a target. These trials cannot be categorized as passive or active and as such cannot be used in the analysis. Therefore, to ensure an adequate number of trials in each condition, out of the 150 cards 2/3 of the cards were transitive and 1/3 were intransitive. *Post-hoc* analysis showed that there was an average of 15.41 (SD: 3.316), 16.72 (SD: 3.235), and 12.34 (SD: 4.292) trials in the passive priming, active priming, and baseline conditions respectively. Participants were only included if the ratio of active to passive priming trials with each avatar was not significantly different from 1.

Coding and Analysis

Priming task Responses during the syntactic priming task were manually coded by the experimenter as being either active or passive. An active sentence is one where the agent of the action is named first (e.g. *the woman kisses the man*) and were coded as 0; a passive sentence is one where the agent of the action is named last (e.g. *the man is kissed by the woman*) and were coded as 1. An independent coder blind to the purpose of the experiment verified that the coding of a random sample of participants was done correctly. Target responses were included in the analysis only if 1) both actors and the verb were named correctly (as a sentence naming only one of the actors does not qualify as a transitive sentence) and 2) no unnecessary information was included in the description (this constrains the participants to using either an active or passive description). We excluded 0.53% (79 out of 14851) of the target responses because they were incorrect.

The responses were analyzed using a mixed-effects logit model, using the glmer function of the lme4 package (versions 1.1-4; Bates et al., 2012) in R (R Core Development Team, 2011). The dependent measure was a binary variable coding whether the response syntax was active (0) or passive (1). The repeated-measures nature of the data was modelled by including a per-participant and per-item random adjustment to the fixed intercept (“random intercept”), with random adjustments to the fixed effects (“random slopes”) as was supported by the data, namely, a random slope *for Prime Type* by participant. We began with a full model and then performed a step-wise “best-path” reduction procedure, removing interactions before main effects, to locate the simplest model that did not differ significantly from the full model in terms of variance explained.

Questionnaire As each participant filled in the Interpersonal Distance Questionnaire thrice (once for each avatar), we conducted a multivariate exploratory factor analysis on these results. However, as this study only consists of a maximum of 198 data entries for the Interpersonal Distance Questionnaire, a value that is too low to conduct an accurate analysis, we combined our questionnaire results with that of similar studies that have used the exact same questionnaire in the same situations (Schoot et al., 2014, 2016; Chapter 3 this thesis). This boosts the total data set to 694 (Kaiser-Meyer-Olkin Measure of Sampling Adequacy: 0.78; Bartlett’s test of sphericity: χ^2 (21): 964.64, p <

.001). Note that this dataset is only used to determine the factor loadings for each question; the factor scores included in the final analysis are only those obtained from the current study. We used parallel analysis to determine the number of factors to be returned by factor analysis. The analysis indicated that 2 factors had the greatest explanatory power for the rating data. Table 4.1 shows the loading values for each of the 2 factors.

Table 4.1 Factor loadings for the Interpersonal Distance Questionnaire. Loadings greater than |0.4| are in bold as these items contribute most to the meaning of a factor. Loadings less than |0.1| are omitted for clarity.

	Factor 1 *Interpersonal Similarity*	**Factor 2** *Shyness*
I could be friends with my partner	**0.75**	-0.19
My partner is similar to me	**0.67**	0.12
My partner appeared generous	**0.51**	-0.32
My partner appeared intelligent	**0.56**	
My partner appeared selfish	**0.51**	-0.13
My partner appeared shy	-0.20	**0.92**
My partner appeared enthusiastic		-0.22
Proportion Explained	0.27	0.15

Results

Avatar Ratings

The participants were asked to rate each avatar on humanness, strangeness, quality of facial expression, and quality of voice in relation to the other avatars to ensure that the avatars were rated the same, despite the new participant group, as they were in Experiment 1. We again found a significant effect of humanness (F = 10.668, p = < .001, Figure 4.2), with Avatar 1 again being rated as the least human. However, the order of the most human and intermediate avatar was reversed in this replication: The most human avatar was now the intermediate avatar and vice versa. The relation of this new intermediate avatar to the other two is still as it was in the previous experiment: It is not significantly different from either the most or least human avatar.

We also found a significant difference in quality of facial expression (F = 12.208, p = < .001) and a trend in strangeness ratings (F = 2.548, p = .081). The ratings for voice were not significantly different between avatars (F = .174, p = .840), but, as in Experiment 1, this was used as a control measure as the exact same voice was used for all three avatars.

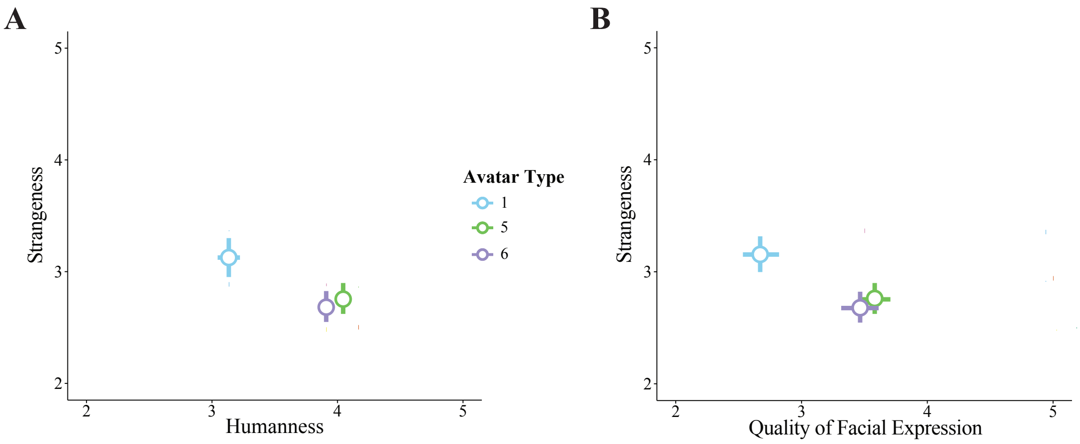

Figure 4.2 Rating of avatar versions for Experiment 2. A. Correlation between the strangeness and humanness ratings, and **B.** Correlation between the strangeness and quality of facial expressions for the six avatars. Error bars represent standard error.

Influences on Syntactic Priming

We assessed syntactic priming behaviour using a logit mixed model. We began with a full model, including avatar type and the two extracted questionnaire factors as well as interactions of all of these with *Prime Type*, and then performed a step-wise "best-path" reduction procedure, removing interactions before main effects, to locate the simplest model that did not differ significantly from the full model in terms of variance explained (Full = AIC: 4115.0, BIC: 4398.2; Best = AIC: 4077.8, BIC: 4190.6; p = .993). Multicollinearity was low (VIF < 2.5). *Prime Type* and *Avatar* factors were dummy coded (all levels compared to a reference group; in this case baseline and avatar 1), *Gender* was sum contrast coded. Continuous predictors were centered.

The fixed effects of the model fit for these data are summarized in Table 4.2.

Table 4.2 Summary of the fixed effects in the mixed logit model for the response choices based on prime structure per avatar type.

Predictor	coefficient	*SE*	*Wald Z*	*p*	
Intercept	-3.84	0.28	-13.90	< .001	***
Gender	0.47	0.28	1.66	.100	
Active Prime	-0.66	0.24	-2.78	.001	**
Passive Prime	1.23	0.19	6.33	< .001	***
Avatar 5	0.00	0.12	0.04	.967	
Avatar 6	0.03	0.12	0.27	.786	
Cumulative Passive Proportion	0.62	0.44	8.21	< .001	***
Avatar 5 * C. Pass. Prop.	-0.75	0.53	-1.43	.154	
Avatar 6 * C. Pass. Prop.	0.38	0.55	0.69	.489	

N= 8777, log-likelihood = -2022.7   * < .05   ** < .01   *** < .001

This model shows a significant influence of passive primes on passive production ($p < .001$) and a significant influence of active primes on active production ($p = .001$). Therefore, we do see a robust priming effect in this experiment. There is also a significant influence of *Cumulative Passive Proportion* on passive production. This factor was calculated as the proportion of passives out of the total transitive responses produced on the target trials before the current target trial. A positive and significant *Cumulative Passive Proportion* therefore suggests that the proportion of passives previously produced positively influences the probability of producing a passive on the current target trial and is commonly used to model the learning effect of priming (Chang et al., 2006; Jaeger and Snider, 2008; Chapter 3 this thesis).

The model, however, does not show an effect of *Avatar* or any influence of any factor from the Interpersonal Distance Questionnaire on priming behaviour (Figure 4.3).

Figure 4.3 shows that there is a trend for a difference in priming magnitude between the different avatar types, however the error bars are quite wide, even with 66 participants, which could explain the lack of statistical significance. One explanation for this is that perhaps analysing priming magnitude per avatar is not the optimal way to analyse this data. Figure 4.4 shows the individual ratings per avatar and it is clear that not every participant rated the avatars the same, even if the average of the ratings are significantly different (as illustrated in Figure 4.2). For example, some participants found avatar 1 much more human than avatar 6, even if the average of the participant group shows the reverse.

Therefore, we re-ran the mixed effects model but instead of avatar type we included each participant's ratings for humanness, strangeness, and quality of facial expression. We began with a full model, including the two extracted questionnaire factors and the three evaluation ratings (excluding quality of voice as this was a sanity check factor) as well as three-way interactions of one questionnaire factor with one evaluative rating with *Prime Type* (e.g., *Shyness*Humanness*Prime*). We also included a quadratic term for humanness and strangeness *post hoc* as figures of the interaction suggested a quadratic term might better fit the data. Random slopes included *Prime Type, Interpersonal Similarity,* and *Strangeness (quadratic)* for Subject, and no random slopes for item. We then performed a step-wise "best-path" reduction procedure to locate the simplest model that did not differ significantly from the full model in terms of variance explained (Full = AIC: 4034.9, BIC: 4360.0; Best = AIC: 4028.6, BIC: 4247.7; $p = .070$). Multicollinearity was acceptable (VIF < 3.96). The results of the best fit model are summarized in Table 4.3.

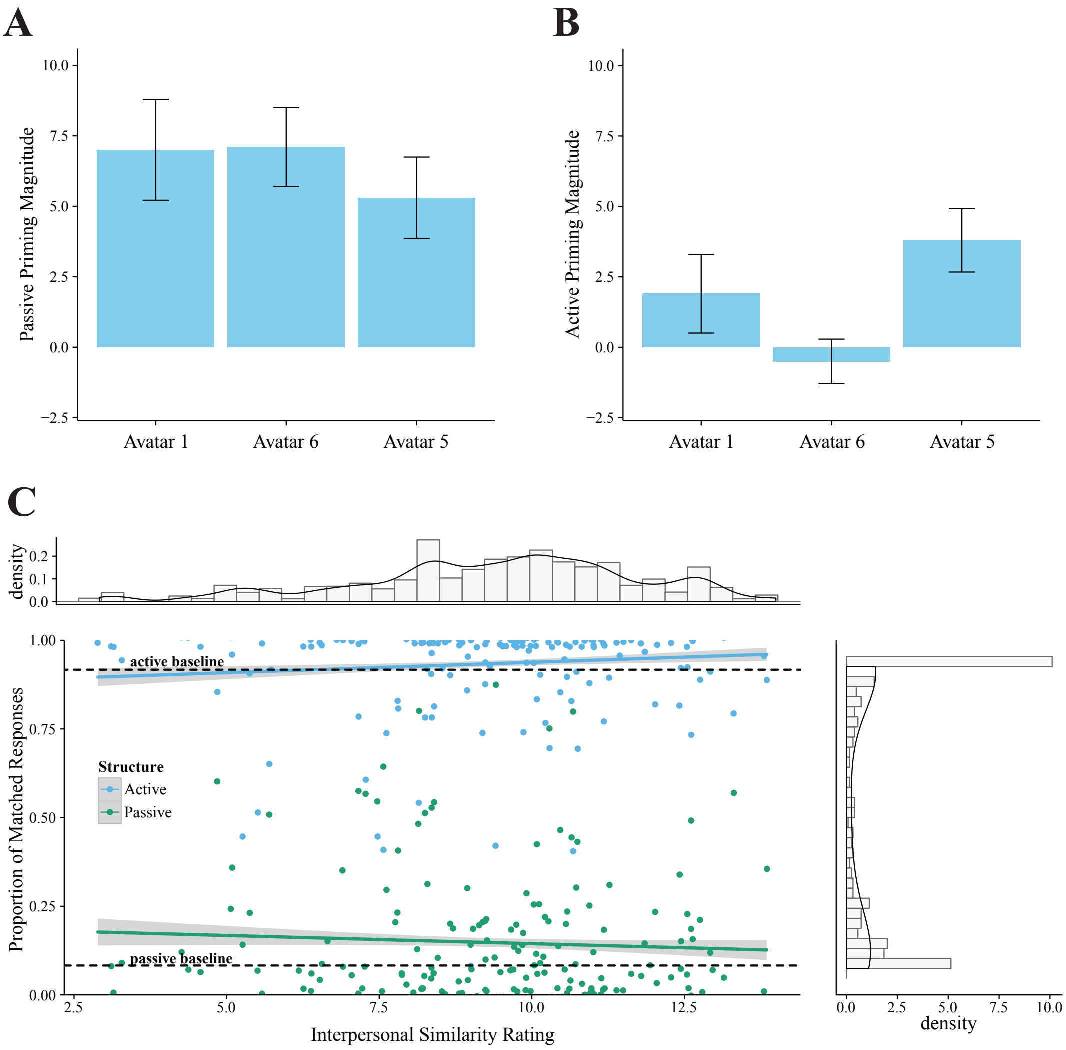

Figure 4.3 Influences on Priming Magnitude. A. Passive priming magnitude per avatar; although there is a significant priming effect, there was no significant difference between partner types ($p > .05$) **B**. Active priming magnitude per avatar; although there is a significant priming effect, there was no significant difference between partner types ($p > .05$). Error bars represent standard error. **C.** There was no significant effect of Interpersonal Similarity rating on priming magnitude for either structure. Error clouds represent standard error.

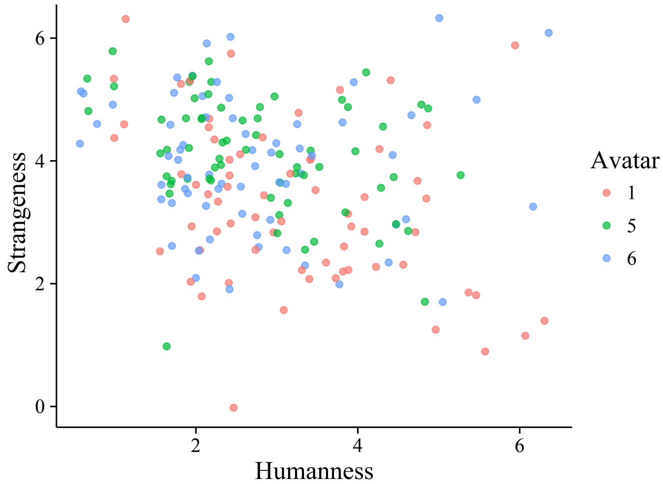

Figure 4.4 Individual ratings per Avatar Shows that even though we used a universal manipulation, individual ratings of the avatars differ dramatically. Therefore analysis will not be done on an avatar-by-avatar basis, but as a function of each participant's individual ratings.

Table 4.3 Summary of the fixed effects in the mixed logit model for the response choices based on prime structure per avatar ratings.

Predictor	coefficient	*SE*	*Wald Z*	*p*	
Intercept	-3.68	0.27	-13.57	< .001	***
Active Prime	-0.77	0.25	-3.12	.002	**
Passive Prime	1.32	0.20	6.60	< .001	***
Interpersonal Similarity	0.34	0.21	1.62	.105	
Strangeness Rating (quadratic)	-0.06	0.08	-0.85	.396	
Shyness	-0.19	0.13	1.51	.131	
Cumulative Passive Proportion	3.40	0.30	11.41	< .001	***
Active Prime * Interpersonal Similarity	-0.10	0.18	-0.56	.574	
Passive Prime * Interpersonal Similarity	-0.27	0.16	-1.66	.098	
Active Prime * Strangeness Rating	0.09	0.08	1.07	.287	
Passive Prime * Strangeness Rating	-0.01	0.07	-0.10	.920	
Interpersonal Similarity * Strangeness	-0.20	0.07	-3.03	.002	**
Strangeness (quad.) * Shyness	-0.15	0.05	-2.93	.003	**
Active Prime * Interpersonal Sim. * Strangeness	0.07	0.06	1.12	.264	
Passive Prime * Interpersonal Sim. * Strangeness	-0.17	0.06	2.84	.004	**

N = 8670, log-likelihood = -1983.3 * < .05 ** < .01 *** < .001

The model again shows priming effects for both active (p = .002) and passive (p < .001) structures, and again an increase in passive production over time (p < .001). For this model, we see influences of the ratings from the Interpersonal Distance Questionnaire and the Avatar Ratings on overall passive production, regardless of prime. This is manifested as interactions between *Interpersonal Similarity* and *Strangeness (quad.)*, and between *Strangeness (quad.)* and *Shyness* which resulted in less passive production as these ratings increase.

In terms of effects on priming, however, we only see a significant three-way

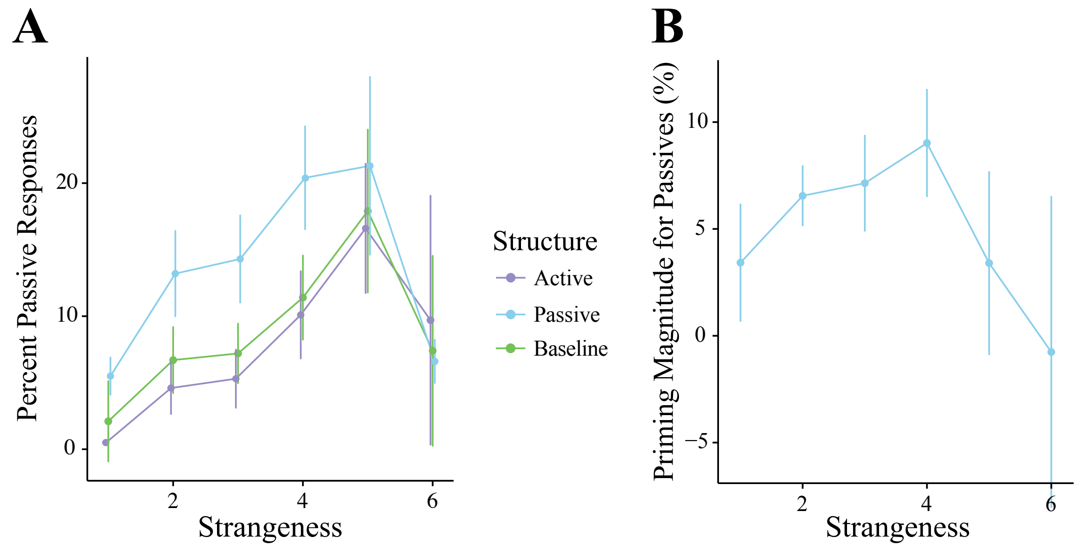

Figure 4.5 Passive Repetition per Strangeness. A. Percent of passive responses after each prime type **B**. Passive priming magnitude (the percent of passive responses after a passive prime after baseline correction) shows an inverted U-shaped curve as a function of Strangeness. Error bars reflect standard error.

interaction between *Interpersonal Similarity*, *Strangeness*, and *Passive Prime*. This three-way interaction is hard to explain, as analyzing the data by splitting the data along either *Interpersonal Similarity* or *Strangeness (quad.)* all provide models without a significant interaction with *Passive Prime* ($p > .135$). We will therefore investigate the contribution of *Interpersonal Similarity* and *Strangeness* on priming separately, as we had planned to before we started the analysis.

The interaction between *Passive Prime* and *Interpersonal Similarity* has been plotted above (Figure 4.3C), which shows a negative, non-significant influence of *Interpersonal Similarity* on passive and active priming magnitude.

Figure 4.5B illustrates how the magnitude of the passive priming effect changes with increasing strangeness ratings: as the strangeness rating increases, the passive priming magnitude increases as well, however, past the (roughly) midpoint of the scale, this upward trend seems to reverse itself. The difference in priming strength between the midpoint and the lowest strangeness rating is roughly tripled (3.42% vs 9.02% passive priming); for the highest strangeness rating this difference is even greater (-0.76% vs 9.02%). This inverted U-shaped curve suggests that at both extremes of our manipulation, participants were least likely to repeat the passive structures of the avatar, whereas in the middle of the manipulation they were most likely to repeat passive structures, an observation that has not been shown before. The active prime condition does not show this relationship and no relationship is seen in relation to the humanness rating of the avatars.

Discussion

In this study we aimed to determine how social factors influence syntactic priming behaviour. As previous studies have reported conflicting conclusions, the direction and strength of influence that social factors have on the magnitude of the syntactic priming effect remained equivocal: do we adapt our language behaviour more or less when we communicate with people we like?

In Experiment 1 we tested 6 avatars to see how human they were, and whether our manipulation (facial expressions) did not cause the avatars to be perceived as unnaturally strange. Previous studies have shown that if the avatar is not human enough, the participants do not prime at all (Beckner et al., 2015; Experiment 2 in Chapter 3 this thesis) and therefore it was imperative to design avatar partners that will still elicit a priming effect that is comparable to human-human interactions. Our results showed that avatars with upper-face animations (in our case, eyebrow movements) were rated as significantly more human than avatars with smiling habits or faster blink rates. These results replicated a trend seen in the agency literature, a field that studies how and why humans perceive that inanimate objects such as dolls and non-humanoid robots have thoughts and feelings of their own. Studies in this field have suggested that the animation and presentation of the eyes and upper face of the being play a major role in determining agency (Looser and Wheatley, 2010), a trend that we were able to replicate here, twice. Even though the differences in our avatars were very subtle, such as an increase in blink completion (close to open interval) from 0.1 seconds to 0.5 seconds or random smile timings versus dialogue-matched smile timings, we were still able to observe strong preferences for some avatars over others. Fortunately, no avatar was rated as uncannily strange (a rating of < 3 out of 6 on the strangeness rating) and therefore we picked the least human, most human, and intermediate out of the 6 avatars tested.

In Experiment 2 we validated our assumption that our avatars were human-like enough to elicit a priming effect. We observed robust priming effects for passives (6.5% increase in passive production after a passive prime compared to baseline) and actives (1.7% increase in active production after an active prime compared to baseline). In line with the human literature, there is a stronger priming effect for passives than for actives, known as the inverse frequency effect (Scheepers, 2003; Ferreira and Bock, 2006; Jaeger and Snider, 2008). We additionally observed a learning effect, such that participants were significantly more likely to produce a passive structure regardless of prime type as the experiment progressed ($p < .001$). The implicit learning nature of structural priming has been discussed before using corpus data (Jaeger and Snider, 2008), as well as observed experimentally (Segaert et al., 2016; Chapter 3 this thesis). Thus we have three arguments to validate our assumption that our avatars were human-like enough to elicit a priming effect akin to those seen in the human literature, with the same characteristics: 1) we observe a robust priming effect, 2) there is an inverse preference effect, i.e. there is a larger priming effect for the less frequent structure, and 3) the priming effect is cumulative over time.

However, the aim of this study was to deduce if there is an interaction between social opinion and language behaviour, and in which direction. We did not observe a significant difference in priming magnitude for each of the three avatars, and we also did not find a significant interaction between priming magnitude and interpersonal similarity with the avatars, regardless of type. We believe the lack of an interaction between interpersonal similarity and priming, which we predicted we would find based on previous literature, could be because the questionnaire was designed for human evaluations. It included questions such as "*I could be friends with my partner*" and "*I find my partner similar to myself*" which may not translate the same when evaluating these statements about a computer. For the priming magnitude per avatar type, we believe the lack of a significant interaction is because by fitting the ratings into three conditions, we may have averaged out the individual opinions of the participants. As we are interested in how the opinion a participant has of a certain avatar influences that participant's language behaviour, it does not matter whether this avatar was previously labelled as the best or least human. For this reason, we reanalyzed the data per avatar rating (humanness, strangeness, and quality of facial expression) to better represent the individual impressions each participant had.

Analyzing the data per ratings indicated that as the participants found the avatars increasingly strange (as an interaction with decreasing *Interpersonal Similarity* and *Shyness*) they produced more passives regardless of prime type. Studies have suggested that the learning part of priming is supported by surprisal (Jaeger and Snider, 2008, 2013). Perhaps the cumulative effect of using an infrequent structure and the strangeness of the partner increased the surprisal effect such that learning increased. In our analysis we had modelled learning as *Cumulative Passive Proportion* and we did not find an interaction of this factor with *Strangeness*, which argues against this interpretation. What drives this increased passive production with increasing strangeness and decreasing values in *Interpersonal Similarity* and *Shyness* is as of yet unclear.

Our novel, and most important, finding for this study is an inverted U-shaped interaction between the strangeness rating of the avatars and passive priming. This interaction was not observed for active priming nor for any of the other avatar ratings. This significant effect was found with the first 36 participants (p = .0004), however, as this inverted U-shaped curve was neither expected nor predicted, we recorded an extra 30 participants. These 30 extra participants also showed an inverted U-shaped curve: a low priming magnitude for the lower and higher ratings (1.28% and 4.36%) and a high priming magnitude for the middle ratings (7.79%). Thus we have such already, albeit indirectly, replicated the finding at least once and are confident in the stability of these results. Our claim for this interaction is based on a three-way interaction between *Interpersonal Similarity*, *Passive Prime*, and *Strangeness* and not a direct two-way interaction. However, we found no significant results when exploring this three-way interaction via median splits of *Strangeness* and *Interpersonal Similarity* and hence we are not entirely sure where this significant effect emerges from. There was a significant correlation between *Interpersonal Similarity* and *Strangeness* which may drive the three-way interaction. However, we have no direct evidence to support this theory and hence further studies are definitely needed to further investigate this effect.

We believe that what we labelled as 'strangeness' when interacting with an avatar interlocutor, as opposed to a human interlocutor, gauges favourability towards the conversation partner. It seems that there is a 'happy medium', in our case a strangeness rating of 4 out of 6, which elicits the highest passive priming effect compared to all the other strangeness ratings (9.02%). As the ratings diverge from this middle rating, the priming effect decreases to either 3.42% for the least strange avatar, or -0.76% for the most strange avatar. This is what has been referred to in other fields as the Goldilocks principle: Something must be within an ideal range to exhibit the maximum effect. There are theories suggesting that priming is a default social behaviour and therefore only occurs if there are no top-down cues to override it (Dijksterhuis and Bargh, 2001). It could be the case that in the middle of our manipulation gradient there are no over-ruling social cues, and therefore the participants exhibit the highest probability of passive structure repetition. It is the extremes, the left and right side of the curve, in which top-down cues override their default behaviour and therefore decrease the probability of passive structure repetition.

Although unexpected, our model suggests that there is a correlation between the interpersonal similarity and strangeness rating, such that the less strange the avatar is perceived, the higher the interpersonal similarity score ($p = .002$). This suggests that our results can be compared to those seen in the human literature: Perhaps the difference between the Balcetis and Dale studies, and Weatherholtz and colleagues may not be a matter of social versus asocial context but perhaps the manipulations tested different ends of this inverted U-shaped curve. Perhaps watching your partner insult a third party causes a stronger opinion than having your partner be mean to just you. Therefore, Experiment 2 from the Balcetis and Dale study might occur on the right side of the curve in Figure 4.5B, whereas their Experiment 1 might occur on the left side, hence explaining why they observe a different interaction between social opinion and priming.

Certain studies have already provided suggestions as to why participants prime less on either side of the 'happy medium'. With a very high interpersonal similarity, participants might attempt to show individuality and creativity by not mimicking their partners. In a study where heterosexual mixed-gender participant pairs were invited to complete a syntactic priming task, the likelihood that the males repeated the syntactic structure used by the female was inversely related to the female's level of fertility (Coyle and Kaschak, 2012). The authors explain this behaviour as the need to show creativity: if the males use "novel" syntactic structures in their responses, they exhibit their creativity and therefore their candidacy as a potential mate. On the other hand, if you interact with someone whom you don't find very similar to yourself, you are already individual and hence do not adopt their behavioural tendencies. This leaves only the person in the middle, who is neither too similar nor too different, that elicits the highest priming effect.

In terms of human-computer syntactic priming studies, there has also been a dispute with what is expected. Previous studies by Branigan and colleagues (2003) and Pearson and colleagues (2006) have shown that participants exhibit a higher priming magnitude with computer partners that were perceived as being less human, whereas Beckner and colleagues (2015) and we (Chapter 3) have shown the opposite effect: Less priming with a computer partner compared to human or human-like partners. This difference can be attributed to how the manipulation was conveyed: for Beckner and our studies, the participant interacted directly with computer partners as the partners were either an avatar presented in virtual reality, or robots seated next to the participant. For the Branigan and Pearson studies, however, the manipulations were more subtle: The participants interacted with the computer via a chat program with no video. As in the human-human syntactic priming studies, this difference in manipulation could cause a difference in the placement of the interlocutors on the inverted U-shaped curve, with the Branigan and Pearson studies starting at the very left, and therefore showed an increase in priming tendency with increasing strangeness perception, whereas Beckner and our studies started at the peak, and therefore showed a decrease in priming tendency with increasing strangeness levels.

To summarize, we were able to observe an inverted U-shaped interaction with passive structure repetition and strangeness rating of the interlocutor, a novel observation that helps piece together previously divergent studies. By taking advantage of the flexibility, yet fine control, that VR offers, we were able to establish that there is an effect of social perception on syntactic priming magnitude in a way not possible using traditional methods. Moreover, we were able to establish that the relationship is of a non-linear nature. In this study we focused on positive/negative ratings although we in no way are claiming that opinion is that unidimensional. The interactions between *Interpersonal Similarity*, *Strangenes,s* and *Shyness* clearly support that there are higher complex relations that influence passive production. Other features such as social goals and motivation may also play a role (although see Schoot et al., 2016). However, our study does show that there is accumulating and convincing evidence that syntactic processing is sensitive to high-level interpersonal factors that can modulate the operation of acclaimed automatic mechanisms.

5

Memory encoding of syntactic information relies on domain-general attentional resources. Evidence from dual-task studies.

In this study we investigate what type of attention is used during syntactic processing. Does syntactic processing require domain-general or language-specific attentional resources? We focus on syntactic priming: In this task, participants listen to a sentence that describes a picture (prime sentence), followed by a picture the participants need to describe (target sentence). We measured the proportion of times participants use the syntactic structure they heard in the prime sentence to describe the current target sentence as a measure of syntactic processing. Participants simultaneously conducted a motion-object tracking (MOT) task, a task commonly used to tax attentional resources. We manipulated the number of objects the participant had to track; we thus measured participants' ability to process syntax while their attention is not-, slightly-, or overly-taxed. Performance in the MOT task was significantly worse when conducted as a dual-task compared to as a single task. We observed an inverted U-shaped curve on priming magnitude when conducting the MOT task concurrently with prime sentences, but no effect when conducted with target sentences. Our results illustrate how, during the prime sentence of the syntactic priming task, attention differentially affects syntactic processing whereas during the target sentence there is no effect of attention on language behaviour. We explain these results in terms of the implicit learning process that is necessary to prime.

Adapted from: Heyselaar E, Hagoort P, Segaert K (under review) Memory encoding of syntactic information relies on domain-general attentional resources. Evidence from dual-task studies. *Quarterly Journal of Experimental Psychology.*

Introduction

Although Wundt in 1900 suggested that language requires attention (Wundt, 1900), most studies investigating the relationship between language and attention have only taken place in the last 20 years. As eye movements and attention are tightly coupled (Deubel and Schneider, 1996), eye gaze shifts and fixations are commonly used in language research as a real-time indicator of where the participant is attending at any given time. For example, studies on spoken word planning have shown that speakers tend to gaze at words and pictures until the completion of phonological encoding (e.g., Korvorst, Roelofs, & Levelt, 2006; Meyer, Sleiderink, & Levelt, 1998) and the seminal paper by Altmann and Kamide (1999) showed that listeners fixate on pictures before they are named, suggesting that we predict upcoming words based on the preceding words. Although these studies have provided evidence that language does require attention, it is still an open question as to what *kind* of attention is used.

There are suggestions that there is not one single pool of attentional resources (Wickens, 1980). Instead, dual-task studies have suggested that at least the visual and auditory domains rely on different attentional resources (Wickens, 1984). For example, Alais and colleagues (2006) illustrated nearly no effect on visual discrimination performance when participants performed a concurrent auditory chord and pitch discrimination task; however, performance decreased when the dual tasks were presented in the same modality. For language, there is no clear consensus on which attentional resources are necessary. It is likely that different aspects of language make different demands on attentional resources since central aspects of language are modality independent, and hence could instead tap into a domain-general "central executive" pool of attentional resources.

A core process of language production and comprehension is the processing of syntax. Syntax refers to the rules that assign grammatical roles and build phrase structure. There is no consensus (yet) on the steps involved in processing the syntax of a comprehended word/phrase (Friederici, 2002 vs. Hagoort, 2003). However, both models are based on ERP evidence which have suggested that some aspects, but not all, occur without the use of attention. The automaticity of syntax is supported by the fact that some steps occur very early (100 – 200ms after word onset; e.g. word category assignment), which is too fast for conscious, non-automatic control. Other steps in syntactic analysis occur later (300 – 600ms after word onset; e.g., morphosyntactic assignment) which is a long enough time period to include steps such as allocation of attentional resources in addition to the syntactic processing steps. Although there is an extensive literature investigating how attention is allocated in both single- (Petersen and Posner, 2012) and dual-tasks (Roelofs & Piai, 2011), whether these resources are assigned from a modality-independent central executive or a language-specific resource pool similar to auditory and visual attention is undetermined.

A common method to measure the processing of syntax is via a syntactic priming task (Bock, 1986). In this task, the participants are exposed to frequently and infrequently used grammatical structures (e.g., *the man kisses the woman* vs *the woman is kissed by the man*) and the probability of the participant re-using the infrequent syntactic structure in their own utterances is used as a measurement of syntactic processing. This task has been used to test multiple characteristics of syntactic processing, such as the memory system used (Ferreira, 2008; Chapter 2 this thesis) or how syntax is learned during development (Kidd, 2012). In the current study we aim to determine whether syntactic processing uses domain-general or language-specific resources. We aim to answer this question by using a dual-task paradigm. If the performance of two simultaneously performed tasks are impaired, it suggests that the processing stages of these two tasks overlap to some extent. Hence increasing attention to one task almost always impairs performance on a second task (Kinchla, 1992), if they tap into the same resources. Otherwise there is no effect on secondary task performance. Therefore, by having participants conduct a secondary task during the syntactic priming task, we can manipulate the availability of attentional resources and measure how that affects syntactic processing.

For our concurrently presented task we will use a motion-object tracking (MOT) task (Pylyshyn and Storm, 1988). In this task, participants are presented with a set of identical balls. A subset of these balls are briefly highlighted to indicate to the participants that they need to track the location of these balls during the next phase of the task. The identical balls then move randomly around the screen for a set period of time. When they stop, the participant either has to indicate the location of the balls they were instructed to track or one ball is highlighted and the participant has to indicate whether this ball is part of the set they had to track. This task therefore requires attention throughout the entirety of a single trial (Scholl, 2008). By manipulating the number of balls the participant has to track, one can control the amount of attentional resources available for other, concurrent tasks. The MOT task has hence been used as a tool with which to manipulate attention in dual-task experiments (Allen et al., 2004; Postle et al., 2005; Fougnie and Marois, 2006). The task relies mainly on visuospatial short-term memory and attention switching (Oksama & Hyönä, 2004). Therefore, if there is an effect of doing this task concurrently with the language task, it is an indication that both tap into the same resources, suggesting that syntactic processing requires domain-general resources.

Due to the nature of the syntactic priming task, the prime portion tests language comprehension as the participants are listening to the picture descriptions, whereas the target portion tests language production as the participants are describing the picture. Therefore, we will run two separate experiments, one in which the MOT task is presented concurrently with the prime portion (hereafter named ‘Encoding phase’ as the dual-task is performed when participants encode the syntactic information) and one when it is presented concurrently with the target portion (hereafter named ‘Retrieval phase’ as the dual-task is performed when the participants retrieve the syntactic information). This will make it clearer when determining how attentional resources are used, as it could be that encoding requires more resources than retrieval.

Although syntax is an essential aspect of language, we predict that syntactic processing does require attention, and particularly domain-general resources due to the modality-independent nature of grammar processing. This would be reflected in our results as a decrease in priming magnitude with increasing attentional load. This study addresses the following questions: 1) does syntactic processing use domain-general resources? 2) how does syntactic processing respond to decreased attentional resources, and 3) is the interaction between syntactic processing and MOT task performance different depending on whether the syntactic information is encoded or retrieved?

Materials and Methods

Participants

70 native Dutch speakers gave written informed consent prior to the experiment and were monetarily compensated for their participation. The participants were divided such that 35 participants completed the Encoding phase (10 male, M_{age}: 22.03 years, SD_{age}: 2.864) and the other 35 completed the Retrieval phase (10 male, M_{age}: 20.80 years, SD_{age}: 2.447).

Statistical Power

Statistical power was calculated using simulated priming data produced by the sim.glmm package (Johnson et al., 2015) in R (R Core Development Team, 2011). For our simulated data set we assumed 20 repetitions per condition and 35 subjects. We assumed a 10% increase in passive production following a passive prime compared to baseline condition, as is commonly seen in the literature (Segaert et al., 2011, Chapters 3 and 4 this thesis). With a difference of 6% between low ball load (low taxing of attention) and high ball load (high taxing of attention), our simulated data has a power of 0.878 with a 95% confidence interval of 0.856 - 0.898.

Materials

Syntactic priming task The pictures used in this task have been used elsewhere (Segaert, Menenti, Weber, & Hagoort, 2011). Our stimulus pictures depicted 40 transitive events such as *kissing, helping* or *strangling* with the agent and patient of this action. Each event was depicted by a greyscale photo containing either one pair of adults or one pair of children. There was one male and one female actor in each picture and each event was depicted with each of the two actors serving as the agent. To prevent the forming of strategies, the position of the agent (left or right) was randomized. These pictures were used to elicit transitive sentences; for each picture speakers can either produce an active transitive sentence (e.g. *the woman kisses the man*) or a passive transitive sentence (e.g. *the man is kissed by the woman*).

Filler pictures were used to elicit intransitive sentences. These fillers depicted events such as *running, singing,* or *bowing* using one actor. The actor could be any of the actors used in the transitive stimulus pictures. These intransitive sentences could be used as fillers, but also as a baseline measurement of each participant's grammatical preferences. The intransitive picture would be used in the prime, with a transitive picture in the target, to measure how the participants would describe such sentences without being primed (baseline trial).

Each experimental list contained 20 targets in each of the 6 transitive priming conditions (active and passive prime for each of the 3 loads) and 20 targets in the baseline condition. We randomly chose from the pictures described above the target pictures to appear in the '1 prime' and the target pictures to appear in the '3 primes' condition; from this we generated 3 counterbalanced lists so that across each triplet of experimental lists the same target picture occurred once with one or three baseline primes, once with one or three transitive primes in the active version and once with the same one or three transitive primes but in the passive version (and we repeated this procedure starting from a different random set 23 times, in order to create 72 experimental lists). Within each experimental list, this resulted in 144 transitive descriptions on target pictures, 144 transitive descriptions on prime pictures and 72 intransitive descriptions leading up to a target in the baseline condition.

The intransitive sentences also served as filler sentences in an extra 72 sentences. In total there were thus 432 sentences in the experiment. Over the whole experimental list 66% of the items (288 out of the total of 432 sentences) elicited transitive sentences.

Task and Design

In order to manipulate the number of attentional resources available, participants completed a standard syntactic priming task and a motion-object tracking task (MOT task) simultaneously. Figure 5.1 depicts the order of events. The task was presented on a desktop computer using Presentation software (script available upon request), the recordings were played over headphones. The syntactic priming task used active (*the man kisses the woman*) or passive (*the woman is kissed by the man*) sentences. To aid understanding, we will describe the designs of each task separately and then describe how we combined them.

Syntactic priming task Each trial consisted of a prime (participants listening to a recording) followed by a target (participants describing the picture using the verb provided). As mentioned above, a prime could be an active sentence (*the man kisses the woman*), a passive sentence (*the woman is kissed by the man*), or an intransitive/baseline sentence (*the man jumps*). A priming effect in our task is therefore defined as the proportion of passive sentences produced after hearing a passive prime, compared to the proportion of passive sentences produced after

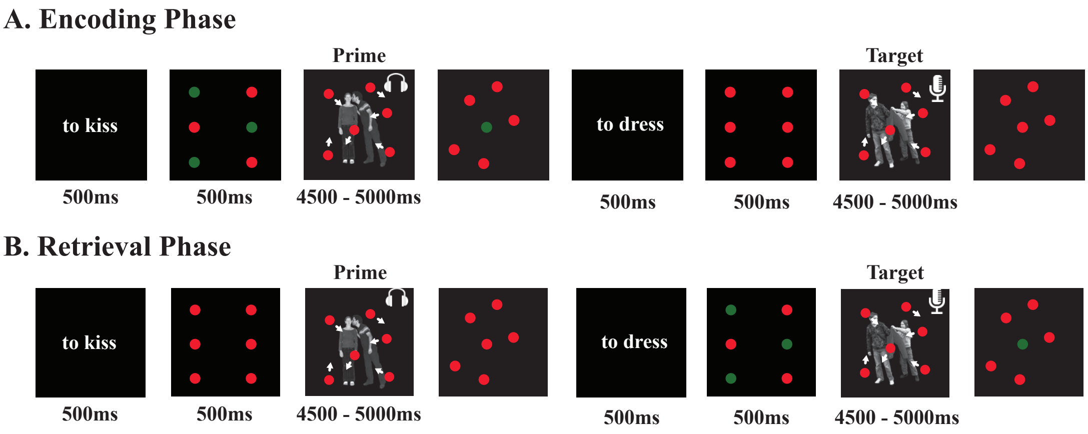

Figure 5.1 Experimental Design Participants completed the dual task either in the Encoding phase (MOT task presented while participants listen to a picture description/ prime phase of the priming task) or in the Retrieval phase (MOT task presented while participants describe a picture/target phase of the priming task). 0, 1 or 3 balls were briefly highlighted at the beginning of the MOT task that the participants have to track. Only one ball is highlighted at the end; participants respond via button press if this was one of the balls they had to track or not. If no balls are highlighted, the participants can effectively ignore the balls for the current trial. Ball load was randomized.

a baseline trial.

Participants were initially presented with a neutral verb (to be used in an upcoming utterance) for 500ms. After 500ms of black screen a greyscale picture would appear. Participants were instructed to either listen to a recording (presented 500ms after picture onset) which describes the picture, or describe the picture themselves using the neutral verb provided earlier. After 4500 – 5000ms (jittered) the picture is removed. The screen is black for an intertrial interval of 1500 - 2000ms (jittered) before the next verb is presented.

MOT task Participants were presented with a 2 by 3 array of 6 identically sized and shaped red balls. A subset of these (none, one or three) were briefly highlighted green for 500ms. After this they would all turn red again and start moving randomly around the screen. After 4500 – 5000ms (jittered) the balls stopped moving. One of the balls is highlighted green and the participant would indicate via key press whether that ball was one of the balls highlighted green at the beginning of the trial or not. If no balls were highlighted at the beginning, then no probe ball is highlighted at the end.

Dual task Each trial began with the presentation of the neutral verb. During the 500ms wait time between verb presentation and picture presentation, the 2 by 3 array of balls would be presented, with the subset highlighted. The picture presentation and the ball movement initiation happened simultaneously to ensure no task started first. After 4500 – 5000ms (jittered) the balls stopped moving and the picture disappeared simultaneously. The intertrial interval of 1500 – 2000ms (jittered) started when the participant responded to the probe ball.

The participants who completed the dual-task in the Encoding phase tracked balls during the prime only (so participants would track balls while listening to picture descriptions) whereas participants who completed the dual-task in the Retrieval phase tracked balls during the target only (participants would track balls while describing pictures). However, in both phases, both prime and target trials involved the presentation of moving balls to ensure the visual input was balanced between phases. No balls would be highlighted at the beginning of the trial, so participants knew they could effectively ignore the balls. The number of balls to track was randomized in one experimental session.

To ensure that participants paid attention to the recordings, 10% of the recordings did not match the picture. The mismatch was balanced between role-switch of the agent and patient, incorrect verb used or incorrect agent/patient used. Participants were instructed to press a certain key if the recording was a mismatch.

Experimental Procedure

Participants were informed that the experiment was about measuring multi-tasking ability. To ensure that the participants understood the task correctly, they first completed practice sessions of the MOT and syntactic priming task separately. The MOT practice session was used to calculate their baseline attentional capacity and contained 10 repetitions of each number of balls to track (0, 1, or 3). The syntactic priming task alone was too short to measure priming magnitude (at least 30 minutes is recommended for a stable effect, Chapter 3 this thesis). No passives were used in the practice session to ensure participants were not primed before the main task began.

At the end of the practice session, participants were able to practice the MOT and syntactic priming task together to ensure they understood the order of events. This contained 5 prime-target trial pairs, of which none were passive structures. During the actual experiment, the participant was given a short, self-timed break every 15 minutes to ensure motivation.

Coding and analysis

Responses during the syntactic priming task were manually coded by the experimenter as either active or passive. Trials in which the descriptions did not match one of the coded structures were discarded. Target responses were included in the analysis only if 1) both actors and the verb were named (a sentence naming only one of the actors does not qualify as a transitive sentence) and 2) the structures used were active, passive or intransitive. In total 43 trials (0.57%) in the Encoding phase and 41 trials (0.55%) in the Retrieval phase were discarded.

The responses were analyzed using a mixed-effect model, using the glmer and lmer functions of the lme4 package (version 1.1-4; Bates, Maechler, & Bolker, 2012) in R (R Core Development Team, 2011). Target responses were coded as 0 for actives and 1 for passives in the factor *Prime*. We used a maximal random-effects structure (Jaeger, 2009; Barr et al., 2013): the repeated-measures nature of the data was modelled by including a per-participant and per-item random adjustment to the fixed intercept ("random intercept"). Our model contained random slopes of *Prime* and *Load*Phase* for the per-participant random intercept, and random slopes of *Load* for the per-item intercept. We began with a full model and then performed a step-wise "best-path" reduction procedure, removing interactions before main effects, to locate the simplest model that did not differ significantly from the full model in terms of variance explained.

Results

Motion Object Tracking Task

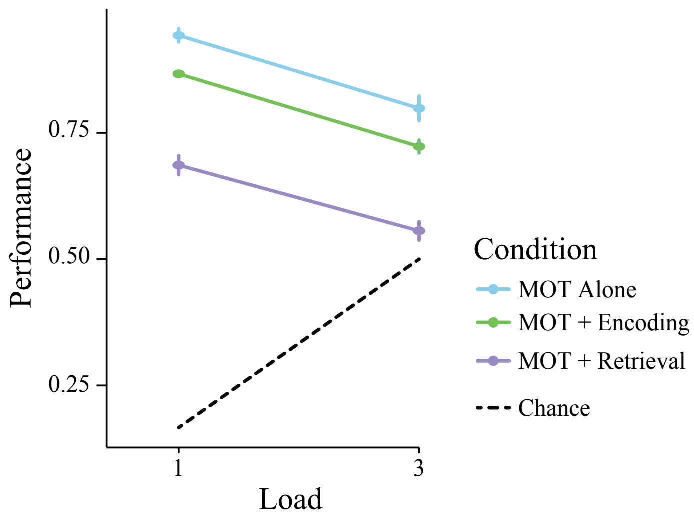

Figure 5.2 Motion Object Tracking Task (MOT) Performance There is a significant difference in the proportion of correct responses between the different conditions compared to performing the MOT task alone. There was a greater drop in performance for the MOT + Retrieval condition than the MOT + Encoding condition. Error bars represent standard error.

Figure 5.2 shows the behavioural results from the Motion Object Tracking (MOT) task. All participants completed the MOT task alone (without a secondary task). Half of the participants additionally completed the MOT task while listening to prime sentence descriptions (MOT+Encoding) and the other half of the participants completed the MOT task while describing the target picture (MOT+Retrieval).

A 3 (MOT Condition: Alone, +Encoding, +Retrieval) x 2 (MOT Load: 1 or 3 balls) between-subjects ANOVA revealed that MOT performance was reduced with increasing loads ($F(1,274)$: 114.13, $p < .001$). More importantly, there was a main effect of condition (F(2,274): 135.85, $p < .001$) showing that performance on the MOT task was significantly reduced. This is consistent with previous dual-task literature: performance of a single task is significantly better than performance of the same task in a dual-task scenario (Bourke, 1996).

Planned comparisons (Tukey's HSD) revealed that, as illustrated in Figure 5.2, both the MOT+Encoding condition and the MOT+Retrieval condition were significantly different from the MOT Alone condition ($p < .001$). Interestingly, performance in the MOT+Retrieval condition was significantly worse compared to the MOT+Encoding condition ($p < .001$).

Syntactic Priming Task

Single task effects Firstly, we performed a logit mixed model on the Load (0) condition, as this is the equivalent of doing the syntactic priming task as a single task. We did this to ensure that our task did elicit a significant priming effect before we investigate whether attentional manipulation influenced the magnitude of this effect. As predicted, there is a significant influence of passive prime ($\beta = 0.79$, $p = .006$; 3.67% average between the Encoding and Retrieval Phase). This indicates that participants primed in our experiment. We additionally observed a significant influence of *Cumulative Passive Proportion* on passive target production ($\beta = 7.67$, $p < .001$). *Cumulative Passive Proportion* was calculated as the proportion of passives out of the total transitive responses produced on the target trials before the current target trial. A positive and significant *Cumulative Passive Proportion* therefore suggests that the proportion of passives previously produced positively influences the probability of producing a passive on the current target trial and is commonly used to represent the learning effect of priming (Chang et al., 2006; Jaeger and Snider, 2008; Segaert et al., 2016, Chapters 3 and 4 this thesis). There was no effect of active prime ($\beta = -0.371$, $p = .325$; -1.48% on average between the Encoding and Retrieval Phase). We are therefore confident that our task elicits the priming behaviour seen in the literature in the absence of an attentional load manipulation (i.e. reverse preference effect and cumulativity; Jaeger and Snider, 2008, 2013; Reitter et al., 2011; Segaert et al., 2016).

As our task elicited a robust passive priming effect akin to the magnitude seen in other studies, we are now able to investigate whether attentional load influenced the magnitude of this effect in the dual-task conditions.

Dual task effects Both dual-task conditions (Encoding phase and Retrieval phase) contained prime-target pairs. During the prime the participant listened to a description of the picture, while during the target the participant described the picture. The only difference in conditions is that for the Encoding phase participants additionally had to complete the MOT task while listening to the prime picture while for the participants in the Retrieval phase, they completed the MOT task while describing the target picture.

Catch rate To ensure that participants paid attention to the recordings, 10% of the recordings did not match the picture. These recordings were played only during the prime portion of the task and hence during the Encoding phase participants listened to the recordings while simultaneously completing the MOT task. During the Retrieval phase, the participants listened to the recordings in a single-task setting, as the MOT task was only presented in the target phase (when the participant describes the picture).

The catch rate was 95.2% (SD: 7.1%) and 91.1% (SD: 6.8%) for the Encoding and Retrieval phases respectively. Neither catch rate was significantly different from what is expected ($\chi^2(2, N = 653) = 5.01$, $p = .082$ for the Encoding phase, $\chi^2(2, N = 630) = 2.18$, $p = .336$ for the Retrieval phase). False alarm rate was 0.8% (SD: 0.48%) and 0.2% (SD: 0.00%) for the Encoding and Retrieval phases respectively. This indicates that even in a dual-task situation (Encoding phase), the participants still listened to the recordings to the same extent as in the single-task situation (Retrieval phase). The results are illustrated in Figure 5.3A.

As the chi-squared test for the Encoding phase had a p-value of .082, we aimed to see whether the catch rate of the Load (1) condition was significantly higher compared to the Load (0) and Load (3) condition (as is suggested by Figure 5.3A). Indeed this is the case ($\chi^2(1, N = 653) = 4.70$, $p = .030$).

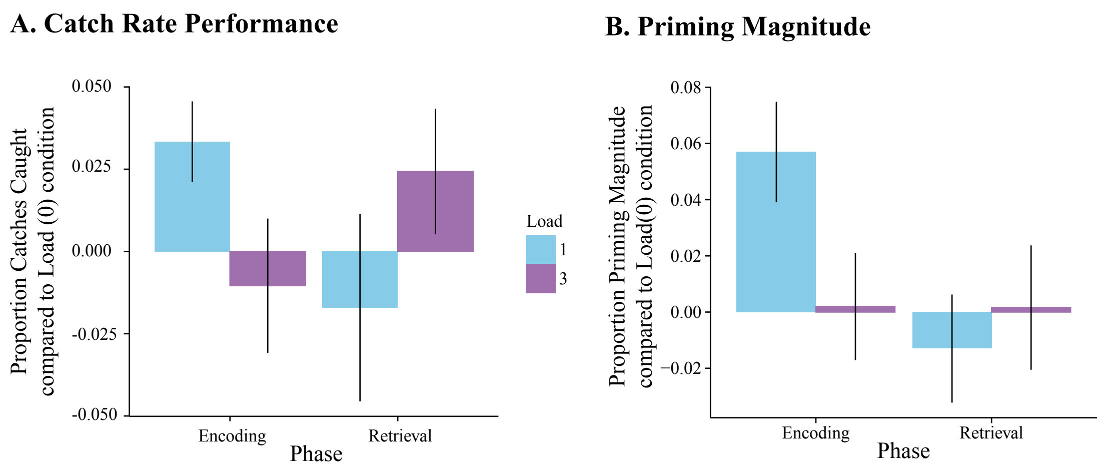

Figure 5.3 A. Catch Rate per phase per load compared to Load (0) condition. The figure illustrates the catch rate performance at each ball load compared to the no ball load (i.e. single-task) condition for each phase respectively. Note that there was no dual-task condition for any of the prime trials in the Retrieval phase. **B. Priming magnitude per phase per load compared to Load (0) condition**. The figure illustrates the amount of priming magnitude difference at each ball load compared to the no ball load (i.e. single-task) condition for each phase respectively. This better illustrates the effect the dual-task scenario has on performance. There is a significant difference in passive priming magnitude between phases ($p = .026$) as well as a Prime by Phase by Load interaction. Error bars represent standard error.

Priming effects Figure 5.3B shows the priming magnitude for each ball load, for each phase, compared to the Load (0) condition. Priming magnitude was calculated as the proportion of passive responses after a passive prime compared to passive responses after an intransitive (not primed) sentence (baseline condition). As stated in *Single Task Effects,* the average priming magnitude for the Load (0) condition was 3.66% passive priming, a low yet robust effect (p = .006). We observed a 7.10% and 4.20% passive priming magnitude for Load (1) for the Encoding and Retrieval Phases respectively, and a 1.60% and 3.90% passive priming magnitude for Load (3) for the Encoding and Retrieval Phases respectively. Figure 5.3B illustrates the difference in priming magnitude at each ball load compared to the no ball (i.e., single-task) condition to better illustrate how the priming magnitude differed compared to the Load (0) condition.

The priming data was analysed using a logit mixed model. We began with a full model (*Prime Type*Load*Phase*Cumulative Passive Proportion),* and then performed a step-wise "best-path" reduction procedure, removing interactions before main effects, to locate the simplest model that did not differ significantly from the full model in terms of variance explained (Full = AIC: 4504.9, BIC: 5105.5; Best = AIC: 4484.4, BIC: 4931.2; p = .430). Multicollinearity was acceptable (VIF < 3.17). *Prime Type* and *Load* were dummy coded (all levels compared to a reference group; in this case baseline and Load (0)). *Phase* was sum contrast coded (each level compared to the mean of the remaining levels, in this case Encoding versus Retrieval). Continuous predictors were centered. The fixed effects of the model fit for these data are summarized in Table 5.1.

Table 5.1 Summary of the fixed effects in the mixed logit model for the response choices based on prime structure.

Predictor	coefficient	*SE*	*Wald Z*	*p*	
Intercept	-3.96	0.23	-17.52	< .001	***
Active Prime	-0.50	0.19	-2.63	.009	**
Passive Prime	0.85	0.17	5.16	< .001	***
Phase	0.48	0.22	2.23	.026	*
Load (1)	-0.00	0.14	-0.03	.975	
Load (3)	-0.10	0.15	-0.64	.520	
Cum. Passive Proportion	4.86	0.68	7.13	< .001	***
Active Prime * Phase	-0.04	0.19	-0.19	.852	
Passive Prime * Phase	-0.53	0.18	-2.95	.003	**
Active Prime * Load (1)	0.23	0.17	1.41	.160	
Passive Prime * Load (1)	0.02	0.14	0.17	.866	
Active Prime * Load (3)	-0.18	0.18	-1.02	.307	
Passive Prime * Load (3)	0.15	0.14	1.08	.279	
Phase * Load (1)	0.19	0.14	1.30	.193	
Phase * Load (3)	-0.25	0.14	-1.75	.080	.
Active Prime * Phase * Load (1)	-0.29	0.17	-1.73	.084	.
Passive Prime * Phase * Load (1)	-0.26	0.14	-1.85	.065	.
Active Prime * Phase * Load (3)	0.34	0.18	1.90	.057	.
Passive Prime * Phase * Load (3)	0.33	0.14	2.36	.018	*
N = 11206, log-likelihood = -2181.2	. < .1	* < .05	** < .01	*** < .001	

The model shows a significant influence of passive primes on passive production ($p < .001$) and a significant influence of active primes on active production ($p = .009$). Therefore, we do see a robust priming effect in this experiment. There is also a significant influence of *Cumulative Passive Proportion* on passive production.

In terms of the aims of this study, there are three-way interactions between *Prime Type* (Active or Passive), *Phase* (Encoding or Retrieval) and *Load* (0, 1 or 3 balls tracked). *Passive Prime* by *Encoding Phase* by *Load (3)* was significant ($p = .018$). To better understand the nature of these three-way interactions, we reanalyzed the data per condition using logit mixed models. We again started with the Full Model described above (without the inclusion of *Phase*) and again used a backward step reduction procedure. Both Best Models were not significantly different from their respective Full Models ($p > .870$). The results of the Best Models are summarized in Table 5.2.

Table 5.2 Summary of the fixed effects in the mixed logit model for the response choices based on prime structure and load.

A. For the Encoding phase

Predictor	coefficient	*SE*	*Wald Z*	*p*	
Intercept	-3.47	3.21	-10.08	< .001	***
Active Prime	-0.42	0.31	-1.35	.178	
Passive Prime	0.14	0.31	0.46	.649	
Load (1)	-0.57	0.33	-1.73	.083	.
Load (3)	-0.10	0.31	-0.32	.750	
Cumulative Passive Proportion	3.65	0.85	4.31	< .001	***
Active Prime * Load (1)	0.29	0.36	0.63	.528	
Passive Prime * Load (1)	0.66	0.33	2.04	.041	*
Active Prime * Load (3)	-0.00	0.34	0.00	.999	
Passive Prime * Load (3)	-0.00	0.32	-0.01	.990	

N = 5490, log-likelihood = -1242.1 . < .1 * < .05 ** < .01 *** < .001

B. For the Retrieval phase

Predictor	coefficient	*SE*	*Wald Z*	*p*	
Intercept	-4.68	0.43	-10.84	< .001	***
Active Prime	0.15	0.53	0.28	.779	
Passive Prime	1.67	0.42	4.00	< .001	***
Load (1)	0.62	0.45	1.37	.169	
Load (3)	0.35	0.46	0.75	.454	
Active Prime * Load (1)	-1.00	0.50	-1.97	.048	*
Passive Prime * Load (1)	-0.46	0.39	-1.16	.245	
Active Prime * Load (3)	-0.49	0.49	-1.00	.319	
Passive Prime * Load (3)	-0.36	0.40	-0.86	.391	
Cumulative Passive Proportion	10.97	1.86	5.91	< .001	***
Active Prime * C. Pass. Prop.	-4.49	2.46	-1.82	0.68	.
Passive Prime * C. Pass. Prop.	-5.22	2.42	-2.16	.031	*

N = 5716, log-likelihood = -941.1 . < .1 * < .05 ** < .01 *** < .001

Table 5.2A and B show that the three-way interaction from Table 5.1 is driven by a significant effect of holding one ball in attention during the Encoding phase on passive priming magnitude (p = .041). This effect is not seen for active priming (p = .528) and there was no interaction between *Prime Type* and *Load* for the Retrieval phase.

This is similar to what was seen for the Catch Rate performance, suggesting that there is a boost in performance in both memory for the syntactic structure (as illustrated by the increase in priming magnitude) as well as the integration of audio and visual streams (as illustrated by the increased catch rate).

Syntactic Priming and MOT

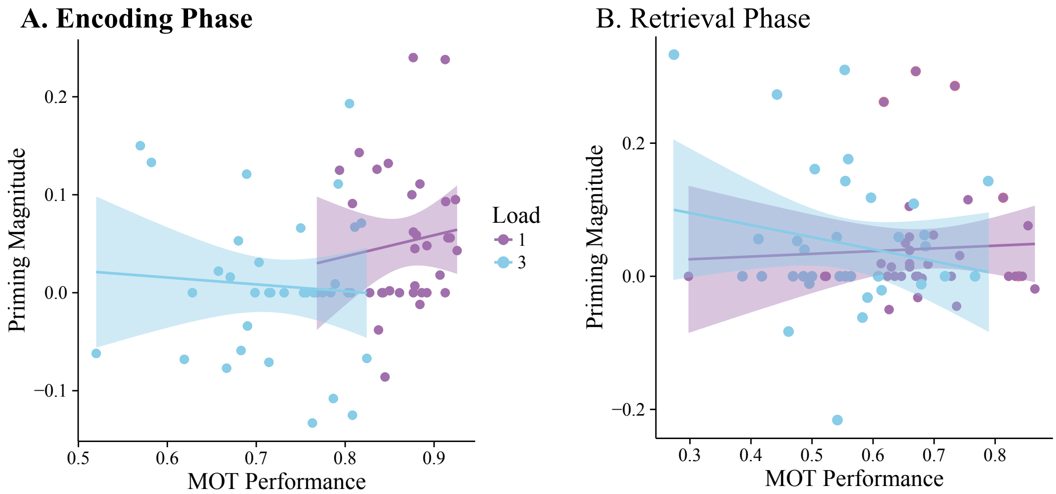

Figure 5.4 Predictability of priming magnitude based on single task MOT performance. The lack of a correlation between priming magnitude and MOT task performance in either the Encoding or Retrieval phases suggests that being good at one task does not predict performance in another task.

We correlated the task performance in the single MOT task condition with priming magnitude, to determine if being good at one task predicts individual performance in the other task. A correlation would suggest that the relationship we have found between the tasks may not be due to shared attentional resources, but due to the fact that some individuals are better at attention-dependent processes, such as goal maintenance and/or persistence.

We show no correlation between the task performance for the Encoding phase (Spearman's rho = 0.079, p = .521) nor for the Retrieval phase (Spearman's rho = -0.093, p = 442). Therefore we are confident that the interaction we see is truly because they tap into the same resources.

Discussion

We utilized a dual-task experiment to determine whether syntactic processing and analysis required language-specific resources or whether it required domain-general resources. To measure syntactic processing we used a syntactic priming paradigm. We modulated the amount of attentional resources available by using a motion object tracking (MOT) task, which is commonly used in the literature in this context (Allen et al., 2004; Postle et al., 2005; Fougnie and Marois, 2006). In addition to modulating attention, we also ran two separate experiments: one in which the MOT task was performed concurrently to the encoding of syntactic information (Encoding phase), and one in which the MOT task was performed concurrently with the retrieval of previously processed syntactic information (Retrieval phase). This will determine whether these two phases use similar attentional resources.

Accuracy in the MOT task was significantly reduced in the dual-task condition compared to the single task condition. This is consistent with the claim that the MOT and the Encoding/Retrieval phases of the syntactic priming task tap into the same resources (Kinchla, 1992). Interestingly, a drop in syntactic priming magnitude was not seen for conditions in which the participants had to track one or three balls compared to conditions in which they had to track no balls. As the MOT task was always presented first, this suggests that the language task was given the priority when it came to resource allocation (Lavie and Tsal, 1994), even though we never told participants to focus more on the language task compared to the MOT task. This result is at odds with another dual-task study that looked at language production and comprehension while driving and found that driving was given the priority over language (Kubose et al., 2006; Bock et al., 2007). However, the authors explained this as driving being given the priority due to the life-threatening nature if it wasn't. This, together with our results, suggests that the natural preference of one task over another is highly sensitive to context.

We next turn to the priming magnitude to determine whether modulation of the availability of attentional resources affects priming magnitude. If we do see a modulation in the priming magnitude, that suggests that the MOT task and syntactic processing share the same resources. If this were not the case, if indeed the attentional resources used by the MOT task were used in some other part of the syntactic priming task that does not involve processing syntax (viewing of pictures, picture comparison to auditory input, etc.) then we would not expect to see a difference in syntactic priming magnitude across the different loads. This is indeed what we found for the Retrieval Phase: although we see a significant drop in performance in the MOT task when conducted simultaneously with the Retrieval Phase, we see no such effect on priming magnitude. We did, however, see a robust *increase* in priming magnitude in the Encoding phase for Load (1). This increase was more than double the priming magnitude observed in the Load (0) condition (4.6% versus 10.4%). This suggests that the Encoding Phase, albeit not the Retrieval Phase, uses the same attentional resources as the MOT task.

This enhancement is not a result we predicted to find. Although *post hoc* and speculative, we find that our result is consistent with a phenomenon known in the field of attention research as the attentional boost (Swallow and Jiang, 2011). When a target appears, no matter if it is a frequent target or not (Makovski et al., 2011; Swallow and Jiang, 2012) attention to this target leads to widespread increases in perceptual processing. The attentional boost hence suggests that there are resources left over in reserves that are allocated when the target appears (dual-task interaction model; Swallow & Jiang, 2013). This is consistent with the results we observe in our current study: participants have assigned the language task as the goal-relevant task and have assigned the majority of their resources to it. However, the appearance of a ball to track causes the participants to increase their perceptual processing while they are trying to keep track of this one ball. This causes them to encode the picture and the auditory input better compared to conditions in which they have no balls to track. When the participants have three balls to track, however, the MOT task already demands extra resources as now the participants have to encode three balls, not one. This means that there are no extra resources to recruit for the boost they would have otherwise received. This enhancement also doesn't occur for the retrieval phase, because they are retrieving stored information, not perceiving anything when they are conducting the MOT task. This does not rule out that the participants in the Retrieval phase do not have a boost during the MOT task, but there is nothing to perceive except the picture as the participants in the Retrieval phase only conduct the MOT task when they are describing the picture. This explanation provides an interesting basis for further research into language and the attentional boost.

Moreover, this increase, or boost, was not only seen in the priming magnitude of the Load (1) condition in the Encoding phase, but we also saw an increase in the catch rate for the same condition, providing converging evidence. This suggests that the increase in priming magnitude is not only limited to enhanced memory for the syntactic structure, but could also be an enhancement in the integration of syntactic structure and visual information. The effect of attention on integration has been observed before in the McGurk effect (McGurk and MacDonald, 1976). The illusion is driven by the integration of audio and visual streams, yet under high attentional load this illusion breaks down, as the integration is not possible with such limited attentional resources (Alsius et al., 2005).

Our results are also interesting in relation to a more general application: multi-tasking while driving, as driving also involves constant spatial attention similar to the MOT task. Previous studies on language and driving have shown that it is not the handling of a cellphone that is dangerous while driving, it is the act of conversing itself (see Strayer, Watson, & Drews, 2011 for a review). Research on memory for language has shown that recall accuracy for a recently comprehended short story is significantly impaired if done while driving (Kubose et al., 2006; Bock et al., 2007). However, if the participant recalls the story while not driving (although it was still told when the participant was driving), there is no significant difference compared to when the participant heard the story while not driving. This result is inconsistent with our study as it suggests that it is not the comprehension/encoding of information that is affected, it is the retrieval of that information. However, many aspects of the Kubose study have been explained as driving being a highly practiced and therefore semi-automatic process. Therefore, perhaps our results are a better reflection of beginner drivers where the task of driving is not as highly practiced.

Overall our results show that language gets priority in terms of assignment of the available resources when it is shared with a non-language (perceptual) task. Although the MOT task was selected due to its “pure” attentional manipulation, it is still a perceptual task, and hence does tap into domain-specific resources, such as visuospatial attention, in addition to domain-general resources. Therefore, our results could also be interpreted as evidence that language taps into visuospatial attention, and hence the interference with the MOT task. There have been suggestions that when people read or listen, they create spatial references in their mind (Fincher-Kiefer, 2001; Fincher-Kiefer and D’Agostino, 2004). However, as we give the participants the picture of the description, we are hard-pressed to suggest that our participants additionally created unnecessary spatial references which could have interfered with the MOT task. Hence we do not believe that the task interference was caused by shared visuospatial resources, but rather shared domain-general/central executive resources.

In summary, our results suggest that syntactic processing does require attention to operate. Even though we do not see a drop in priming magnitude in the language task, we do see a drop in the MOT task performance, meaning that the MOT task had less resources to complete the task accurately. This could only have occurred if another task was tapping into the same pool of resources. It also suggests that language receives priority in terms of assignment of the available resources when it is shared with a non-language (perceptual) task. Additionally, a modulation in priming magnitude during the Encoding Phase caused by manipulating the number of attentional resources available suggests that syntactic processing uses the same attentional resources as the MOT task. The attentional boost effect seen in the Load (1) Encoding phase condition is interesting and has not been observed before for modality independent processes. It poses the question if this effect can be seen for other non-automatic language processes and what role this effect could play in language comprehension.

6

Changes in alpha activity reveal that social opinion modulates attention allocation during face processing.

Conducting a task with a partner, compared to doing the task alone, influences participants' performance in the task. Additionally, performance and behaviour of the participant has been shown to vary as the opinion they have of their partner shifts from more to less favourable. In this study we investigate whether allocation of attention could explain some of these effects. By looking at the modulation of power of alpha oscillations as participants viewed pictures of partners they had previously either interacted with or not interacted with, we found that processing the face of a person with whom a rapport was developed captured more attention. Additionally, we observed a U-shaped change in alpha power as a function of evaluative ratings, suggesting that participants pay most attention when the partner is neither too favourable nor least favourable. Our study therefore provides evidence that attention plays a role in how we perceive interaction partners.

Adapted from: Heyselaar E, Mazaheri A, Hagoort P, Segaert K (in prep) Changes in alpha activity reveal that social opinion modulates attention during face processing.

Introduction

For over 100 years research has shown that an individual's task performance when interacting with another person is different compared to the performance when the individual does the task alone (Triplett, 1898; Burnham, 1910). This effect is seen in behaviours such as eating (Herman, 2015), cognition (Bond and Titus, 1983), and even when participants are interacting with human-like computers (Mandell et al., 2015; Qu et al., 2015). It is commonly referred to as the social facilitation effect.

While there have been many theories postulated to account for this phenomenon, for the purposes of this article we will focus on those involving the allocation of attention, of which the two most prominent are the distraction-conflict hypothesis and the feedback-loop model. The distraction-conflict hypothesis (Sanders, Baron & Moore, 1978; Baron, 1986) proposes that the presence of others is a distraction, which leads to attentional conflict in terms of cognitive overload and selective focusing of attention. Thus here the presence of others always leads to impeded performance on difficult tasks, while it is the number of distractors in the environment that determines whether there is a change in performance on simple tasks. On the other hand, the feedback-loop model suggests that the presence of an audience causes participants to focus attention on the self (Duval and Wicklund, 1973; Wicklund, 1975). This in turn leads to thoughts about discrepancies between the actual self and the ideal self. According to this theory, participants thus improve their performance in the presence of others because of this increased awareness about their own behaviour. There are many more theoretical proposals, each with a different hypothesis on how attentional allocation can explain behavioural effects (for a recent review, see Strauss, 2002). Interestingly though, there have as of yet been no neuroimaging studies looking at how attention is allocated during a task while interacting with a partner compared to doing the task alone and whether this is indeed playing a factor in the behaviour seen. The first objective of our current study is to find neuronal evidence that attentional resources are differentially allocated when participants viewed faces of partners they have just interacted with or not.

The second objective of the current study is to measure whether the allocation of attention differs as a function of the opinion the participant has of their interaction partner. This objective is motivated by previous literature showing that the opinion the participant has of their partner (in interaction tasks) can also influence their behaviour (Lott and Lott, 1961): people who rated their partner as more likeable conformed their communication and behaviour towards their partner. For the Lott and Lott study, participants changed their opinion to better match what they thought their partner's opinion was if they rated their partner as likeable. For language, this has, for example, also been shown in syntactic priming studies. Syntactic priming refers to the phenomenon in which participants adopt the grammatical preferences of their partner and thus throughout the length of an experimental session, participants increasingly use their partners' grammar in their own utterances (Bock, 1986a). Recent studies have shown that the magnitude with which participants adapt their language behaviour to match their partner varies as a function of how the participant rated their partner. In one study this is seen as a modulation of whether the participant liked or disliked their partner (Balcetis and Dale, 2005), in another it is shown as a function of what the authors called "interpersonal similarity" (Weatherholtz et al., 2014). Although these studies used different measurements of what we will refer to as "social opinion", whether the participant found their partner more or less likeable or found themselves more or less similar to their partner, modulated the participant's behaviour (although see Schoot et al., 2016). We hypothesize that this influence on behaviour could also be mediated by attention allocation. It has been proposed that syntactic priming is mainly supported by implicit learning (Bernolet, Collina, & Hartsuiker, 2016; Chang, Dell, & Bock, 2006; Chapter 2 this thesis): participants unconsciously pick up on the increased use of rarely-used grammatical structures uttered by their partner (Chang et al., 2006) such that it gets incorporated into their own speech. If the participant were distracted by features of their partner, they might be unlikely to encode this information as well, resulting in less integration of those structures into their own speech. Specifically, our second hypothesis is thus that the social opinion the participant has of their interaction partner should modulate how much attention is allocated when interacting with that person. Whether that attention is allocated more to the partner, or more to the self (in line with the feedback-loop model), as a function of social opinion is unclear.

Electroencephalography (EEG), a non-invasive neuroimaging technique that measures the electrical potential generated by neurons, provides a window into the working human brain during cognitive processing. The EEG signal contains oscillatory activity in distinct frequency bands that have been found to map on to different facets of cognition (Siegel et al., 2012). Alpha activity, an oscillation occurring at a frequency of 10 Hz, has been suggested to play a pivotal role in attentional allocation (Foxe et al., 1998; Klimesch et al., 2007; Mazaheri and Jensen, 2010). Specifically, alpha activity has been found to increase in task-irrelevant regions, while being suppressed in task-relevant regions (Yamagishi et al., 2003; Kelly et al., 2006; Rihs et al., 2009; Foxe and Snyder, 2011; Mazaheri et al., 2014). However, these studies have all directed attention using explicit, sensory cues (such as bright flashes, for example). Comparatively very little prior work has been done investigating the examined attentional modulation of alpha activity (or any other oscillation) induced by non-sensory, implicit factors such as opinion.

A modulation in alpha band power as a function of social opinion would suggest a change in attention. In Chapter 4 we used computer-based confederates ("avatars") in virtual reality to show a change in the magnitude of the syntactic priming effect as a function of the ratings of said avatars. Therefore, we will use those same avatars in this study. Although an avatar is viewed as a computer, we believe that any effects we observe in this study will be representative of doing the same experiment with human partners, similar to what we showed in Chapter 3.

In this study we will use the avatars used in previous studies to determine if this rating modulates alpha band power. Additionally, by using the high temporal resolution EEG offers we can also investigate whether this allocation of attention occurs top-down or bottom-up. Studies have suggested that emotional stimuli rapidly attract attention at early stages of sensory analysis (Lang et al., 1997) whereas a later influence of attention on stimuli processing would suggest a top-down influence on behaviour. The first aim of our study is thus to determine whether viewing an avatar the participant has interacted with results in a different degree of attentional allocation (compared to viewing an avatar the participant has not interacted with) as investigated through examining changes in alpha activity. The second aim is to determine whether the amount of attention allocated (measured as modulations in alpha activity) to the avatar varies as a function of the participants' social opinion of that avatar.

Participants will be asked to view pictures of the avatars while we record their brain activity via EEG (Phase 1). This will function as baseline activity for viewing the pictures before the participants interact with the avatars depicted in the pictures. The participants will then interact with each avatar in Virtual Reality (Phase 2). After each interaction, the participant will be asked to evaluate the avatars in three categories, or traits. These three traits (perceived humanness, perceived strangeness, and quality of face) have previously been linked to a modulation in syntactic priming magnitude (Chapter 4 this thesis) such that there was an inverted U-shaped interaction between strangeness rating and priming magnitude. The participants showed the highest syntactic priming magnitude when they interacted with avatars they rated as medially-strange (a rating of 4 out of 6). They showed a significantly lower priming magnitude when interacting with avatars they found less or more strange compared to this middle-rated avatar. If this modulation is caused by a modulation in attention, we expect to see a relationship between alpha power and these ratings in Phase 3. In Phase 3, the participants will view the same pictures they viewed in Phase 1 while we record their brain activity. We will also include one avatar that the participants do not interact with, but will view pictures of in Phases 1 and 3. Therefore, as our first aim is to determine whether viewing a person the participant has interacted with results in a different degree of attentional allocation, we expect to see a difference in alpha activity between the Non-interacted-with avatar and the Interacted-with avatars in terms of alpha band power, namely. This study will therefore provide electrophysiological evidence which may clarify why participant behaviour is different when conducting a task with a partner. The second aim of this study is to determine whether the amount of attention allocated, as gauged through modulation in alpha power, to the Interacted-with avatar varies as a function of the participants' social opinion of the avatar, as measured through the ratings of the three traits discussed earlier. Previous studies only found an effect as a function of perceived strangeness, however, we will include all ratings in our analysis.

Materials and Methods

Participants

30 native Dutch speakers (2 male, M_{age}: 21.53 years, SD_{age}: 2.60) gave written informed consent prior to the experiment and were monetarily compensated for their participation. As the EEG cap had to fit underneath the virtual reality (VR) helmet, we were limited to testing participants with small head sizes (58cm diameter and below), restricting us to mostly female participants.

Experimental Procedure

The participants were informed that there were three phases to the experiment, but only received detailed information about Phase One at the beginning of the experiment. At the start of Phase Two they were informed of the goal of the study. The entire procedure is summarized in Figure 6.1.

Participants initially viewed 560 static photos, of which 400 were of the 4 avatars. The faces of the four avatars were all exactly the same, and hence to be able to discriminate between photos of them, they were given different shirt colours. For Phase Two, participants interacted with three of the four avatars in Virtual Reality. Here the avatars were animated such that each avatar had different facial expressions in terms of their smile habit, eyebrow movements

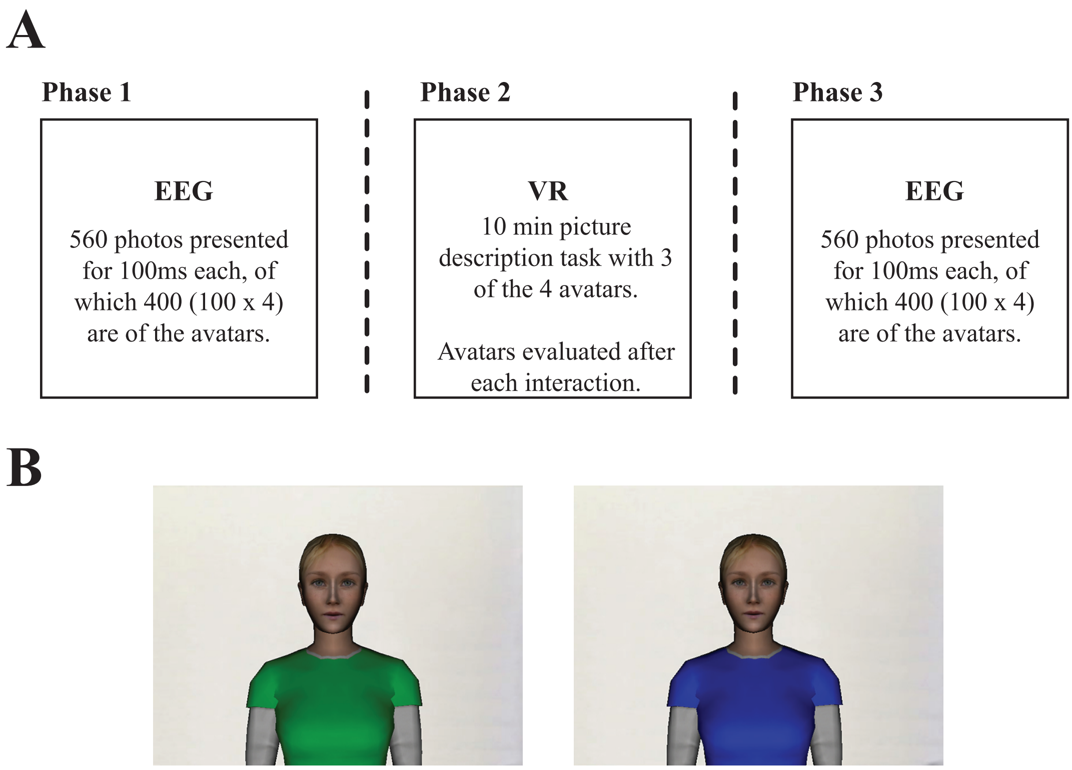

Figure 6.1 A. Procedure B. Examples of two of the four avatars. Avatars were presented with green, blue, red and yellow shirts

and blink rate (see Table 6.1). Participants were therefore able to form opinions about the three avatars based on these facial characteristics. For Phase Three, the participants were again shown static pictures. The participant used the colour of the shirt to discriminate between the avatars they had interacted with (and formed an opinion of) and the ones they had not interacted with. Details of each phase is given below:

Phase 1 – Picture evaluation task In Phase 1 participants viewed photos of avatar and filler pictures, with the aim to measure EEG responses of the participants to the avatar pictures before they had formed any opinion of these avatars. The participants were instructed to evaluate pictures as either "likeable" or "not likeable." Each trial started with the presentation of a fixation cross for a 400 – 600ms jittered interval. This was followed by the presentation of a picture in the center of the screen. Pictures were presented for 100ms, followed by a jittered interval of 2000 – 3000ms before an evaluation screen was presented. Participants were asked to indicate, via a button press, whether they liked or disliked the picture they just saw. The location of the options (left or right) was randomized between participants.

Phase 1 consisted of 560 pictures. 400 of these pictures were of the avatars (100 repetitions of each avatar); the avatars are described below (*Materials*). The remaining 160 were pictures selected from the pictures described below (*Picture Evaluation Task*). Picture order was randomized between participants.

Phase 2 – Picture description task Participants interacted with three avatars for 10 minutes each in the virtual environment (these avatars are hereafter referred to as "Interacted-with avatars"). The EEG cap remained on in Phase 2, but no activity was recorded. The fourth avatar was not interacted with as a control (hereafter "Non-interacted-with avatars"). The shirt colour of the Non-interacted-with Avatar was pseudo-randomized across participants.

Participants at this point were informed of the goal of the study, to ensure they paid as much attention to the relevant characteristics of the avatar as possible.

The participant and the avatar would alternate in describing picture cards to each other. If the listener saw the card described by their partner as one of the cards in their spread they would select it, causing it to be automatically replaced by a novel card. The listener would then become the speaker and pick a card to describe. This continued until 50 cards were described, after which the headset was removed and participants were asked to fill out a pen-and-paper questionnaire about the avatar. We favoured a pen-and-paper questionnaire instead of having the avatar ask the questions directly as previous research has shown that if the participant evaluates the avatar in the presence of said avatar, they rate them more favourably (Nass et al., 1999).

The questionnaire consisted of three 6-point Likert-scale questions asking to rate the avatar on perceived humanness, perceived strangeness, and quality of their facial expressions in relation to the other two avatars. After each avatar, the participants were allowed to change their ratings for previously viewed avatars. The order of the avatars was pseudo randomized across participants.

Phase 3 – Picture evaluation task In Phase 3, we again recorded EEG activity while participants viewed photos of the four avatars mixed with filler pictures. We kept Phase 3 the same as Phase 1, using the same picture list, therefore any difference in the modulation of alpha activity induced by viewing pictures of the avatars between Phase 1 and Phase 3 would like be due to the social interaction that occurred in VR in Phase 2. By comparing the Interacted-with avatar to the Non-interacted-with avatar we can control for any repetition effects as both are viewed in both Phases. Before EEG recordings in Phase 3 began, impedance for each electrode was checked and adjusted as necessary.

Materials

Picture evaluation task (Phase 1 and Phase 3) 165 pictures were taken from the Geneva Affective PicturE Database (GAPED; Dan-glauser & Scherer, 2011). An equal number (55) belonged to the category positive, negative, or neutrally-rated pictures, in terms of emotional valence. We attempted to ensure that the arousal rating was comparable between picture categories as much as possible. The average ratings for arousal were thus 30.47 (SD: 9.491), 29.52 (SD: 5.92), and 46.46 (SD: 7.01) for positive, neutral, and negatively rated pictures respectively.

We also included four avatar photos: the same picture of the avatar with a yellow, green, red, or blue shirt.

Avatars and virtual environment in Phase 2 All avatars had the same exterior adapted from a stock avatar produced by WorldViz (“casual15_f_highpoly”; see Figure 6.1B). All the avatars’ speech was pre-recorded by the same human female and played during appropriate sections of the experiment. The avatars’ appearance suggested that she was a Caucasian female in her mid-twenties, which matched the age and ethnicity of the Dutch speaker who recorded her speech.

The three facial expressions used have been tested in Chapter 4 to represent the least, intermediate, and most human avatar. These three facial expressions involved combinations of subtle changes in blink rate, smiling and eyebrow habits (Table 6.1). Facial expression choices were based on work done by Looser & Wheatley (2010) who have shown that perception of humanness is dependent on upper face movement.

Blinks happened once every 1 - 5 seconds. For versions with normal smiling and normal eyebrow habits we explicitly programmed when the avatar would smile and/or raise her eyebrows, such that it would coincide with the content of her speech. For example, the avatar would raise her eyebrows when asking a question and smile when she was enthusiastic. When not speaking, she would smile once every 5 - 10 seconds and raise her eyebrows once every 1 - 5 seconds such that she would still differ from the no smile/no eyebrow version. All of these changes were extremely subtle to ensure that they can still be related to ecologically valid behavioural characteristics that one would encounter in the everyday world.

Table 6.1 Avatar Facial Expressions

Avatar	Blink Duration[1]	Smiling Habit	Eyebrow Habit
1	No blink	No smile	No movement
2	0.1 seconds (Normal)	Dialogue-matched	Once every(3-5 seconds)
3	0.1 seconds (Normal)	Dialogue-matched	Dialogue-matched

[1]Measured from the beginning of the closing movement to when the eye is fully open again

The virtual environment (VE) was a stock environment produced by WorldViz ("room.wrl") adapted to include a table with a wooden divider. We chose to have the cards displayed at the top of the divider so that the participants could see the cards while facing forward. This was done due to the weight of the head-mounted display (HMD), which would cause an uncomfortable strain on the back of the participants' heads when they face down. Having the participants face forward throughout the entire experiment distributes this weight more comfortably.

The experiment was programmed and run using WorldViz's Vizard software. Participants wore an NVIS nVisor SX60 HMD, which presented the VE at 1280 x 1024 resolution with a 60-degree monocular field of view. Mounted on the HMD was a set of 8 reflective markers linked to a passive infrared DTrack 2 motion tracking system from ART Tracking, the data from which was used to update the participant's viewpoint as she moved her head. It is known that this type of headset can cause dizziness and nausea due to the exclusion of the participant's nose in the field of view (Whittinghill et al., 2015). However, as each interaction was quite short (~5 minutes), none of our participants reported feeling any nausea.

Additionally, a single reflective marker was taped onto the index finger of the participant's dominant hand. This marker was rendered as a white ball in the VE, such that participants knew the position of their finger at all times. Sounds in the VE, including the voice of the avatars, were rendered with a 24-channel WorldViz Ambisonic Auralizer System.

Picture description task in Phase 2 The pictures used in this task have been described elsewhere (Segaert, Menenti, Weber, & Hagoort, 2011)(Menenti, Gierhan, Segaert, & Hagoort, 2011). Our stimulus pictures depicted 40 transitive events such as *kissing, helping,* or *strangling* with the agent and patient of this action. Each event was depicted by a greyscale photo containing either one pair of adults or one pair of children. These pictures were used to elicit transitive sentences; for each picture speakers (avatars and participants) can either produce an active transitive sentence (e.g. *the woman kisses the man*) or a passive transitive sentence (e.g. *the man is kissed by the woman*).

We also included pictures depicting intransitive events such as *running, singing,* or *bowing* using one actor. The actor could be any of the actors used in the transitive stimulus pictures. Each card consisted of one stimulus picture with the relevant verb printed underneath.

Data Analysis Approach

Pre-processing EEG was recorded from 64 cap-mounted Ag/AgCl electrodes (ActiCAP, Brainproducts). Horizontal eye movements were monitored by two electrodes placed at the outer left and right canthi. Vertical eye movements were monitored using an electrode placed below the left eye. In addition, electrodes were placed on the left and right mastoid bones. During EEG recording, all electrodes were referenced to the left mastoid. All impedances were kept below 10kΩ. Signals were recorded with a BrainAmp amplifier system, using a 150 Hz low-pass filter, a time constant of 10s (0.016 Hz), and a 500 Hz sampling rate. Signals were later re-referenced off-line to linked mastoids.

The pre-processing of the data was done using functions from EEGLAB (Delorme and Makeig, 2004) and the Fieldtrip software package (Oostenveld et al., 2011). Fieldtrip EEG epochs were locked to the onset of the picture and manually inspected for non-physiological artefacts. Ocular artefacts were removed using independent component analysis (infomax algorithm) incorporated as the default "runica" function in EEGLAB.

Time-Frequency Representations (TFR) of power Time-frequency representations (TFR) of power were calculated for each trial using sliding Hanning tapers having an adaptive time window of three cycles for each frequency of interest ($\Delta T = 3/f$), utilizing the same approach used in previous studies e.g., Mazaheri et al. (2014) and van Diepen, Cohen, Denys, & Mazaheri (2015).

Measurements of spectral power Our selection of frequency bands and time windows of interest were based on the grand-average data across conditions (all avatars, Phase 1 and Phase 2; seen in Figure 6.2, consistent with the suggestions of Brooks et al. in press). These data were then normalized as percent change from a baseline interval (650 – 150ms before picture onset) within participants to reduce the contribution of participants with large variance in the power estimates. Frequency and time points of interest were then identified as delta (2 – 3.14 Hz) occurring at 140 – 400ms post stimulus, an early alpha (11 – 14 Hz) suppression occurring 250 – 500ms post-stimulus, a later alpha (9 – 14 Hz)

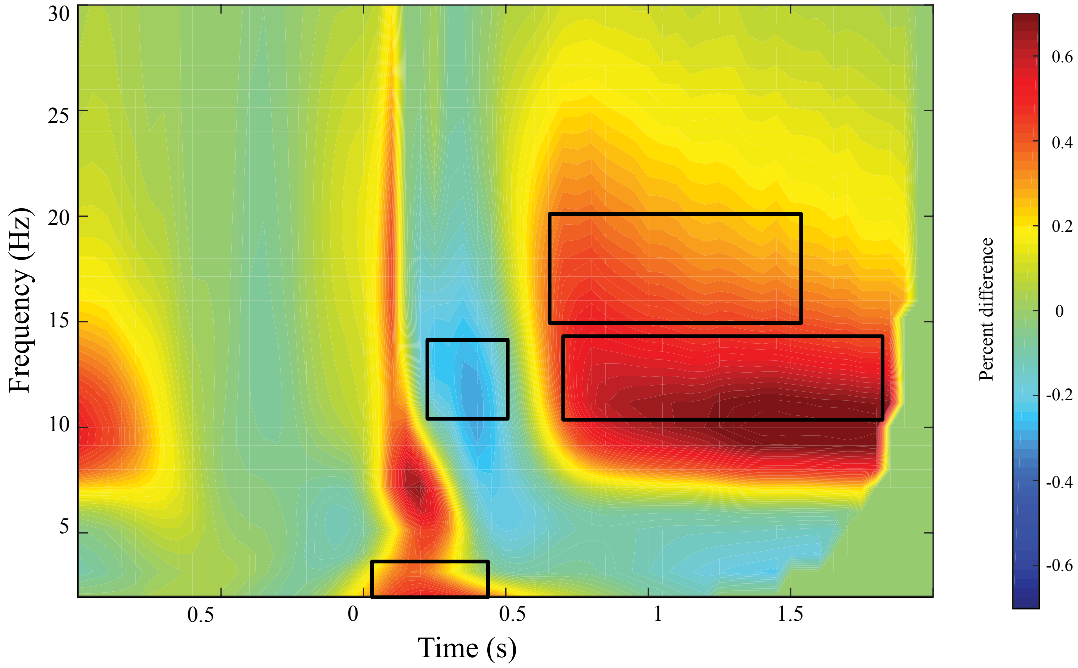

Figure 6.2 TFR of all conditions (all avatars, Phase 1 and Phase 3) for occipital electrodes. We used this TFR to visually identify frequency and time bands of interest (outlined), which were then subjected to a cluster-level test to determine if there is a significant difference between the Interacted-with and Non-interacted-with avatars in these frequency and time regions of interest.

increase occurring 800 - 1800ms post stimulus, and beta (15 – 20 Hz) occurring 700 – 1500ms post stimulus.

To control for low-level sensory-induced oscillatory changes in the EEG induced during the onset of pictures, we subtracted the EEG activity in Phase 1 from the EEG activity in Phase 3. This ensures that we are subtracting any EEG activity related to visual onset of the pictures (and performing the 'like/dislike' task) since this occurred in Phase 1 as well as in Phase 3. These data were then normalized as percent change from the baseline interval (650 – 150ms before picture onset).

Correction for multiple comparisons We corrected for multiple-comparisons (multiple electrodes) by means of a cluster level (over-electrodes) randomization routine (Maris and Oostenveld, 2007). According to this routine Monte Carlo *p* values were obtained by randomly swapping the conditions 1,000 times within subjects and calculating the maximum cluster-level test statistic. Type 1 errors because of multiple comparisons are reduced by clustering neighbouring electrodes that show a similar effect (p = .05, two-sided).

Coding and Analysis

Mixed models The values extracted from the electrodes of interest (see *Results*) were analyzed using a linear mixed effects model, using the lmer function of the lme4 package (versions 1.1.9; Bates, Maechler, & Bolker, 2012) in R (R Core Development Team, 2011). The dependent measure was the values extracted from the regions of interest. The repeated-measures nature of the data was modelled by including a per-participant and per-shirt-colour random adjustment to the fixed intercept ("random intercept"). We began with a full model (two-way interactions between each of the ratings) and then performed a step-wise "best-path" reduction procedure, removing interactions before main effects, to locate the simplest model that did not differ significantly from the full model in terms of variance explained (Full Model = AIC: 282.55; BIC: 316.58; Best Model = AIC: 277.74; BIC: 292.32; p = .191). The best model included main effects of Perceived Strangeness and Perceived Humanness. All ratings were centred before entry into the model. *P* values were extracted using the Anova function from the car package (version 2.1.0; Fox & Weisberg, 2011) using Wald Chi-Square tests (Type III).

Results

The onset of the pictures, irrespective of avatar or condition, induced a suppression in the power of alpha activity peaking at around 400ms, followed by an increase starting at 800ms after the onset of the picture.

I. Interacted-with avatars induced greater alpha suppression than Non-interacted-with avatars

We set out to investigate whether oscillatory activity is modulated by the experience of interacting with and evaluating the traits of (both occurring in Phase 2) 3 of the 4 avatars.

The comparison of alpha power between Interacted-with and Non-interacted-with avatars revealed that for the Interacted-with avatars alpha was significantly more suppressed at 11 - 14 Hz, 400 – 450ms post stimulus onset (Monte Carlo p = .035, corrected for multiple comparisons; Figure 6.3A) and at 9 - 14 Hz, 800 – 1000ms post stimulus onset (Monte Carlo p = .042, corrected for multiple comparisons; Figure 6.3B). There were no significant differences for this comparison in the delta or beta range.

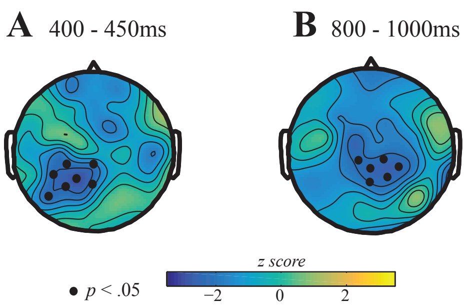

Figure 6.3 Topography of the alpha power of the difference between Interacted-with avatars and Non-interacted-with avatars, after subtracting Phase 1 from Phase 3. This comparison reflects processing related only to having interacted with the avatar and evaluating their traits (both occurring in Phase 2). There were two clusters: **A,** a cluster at 11 – 14 Hz, 400 – 450ms post stimulus onset; and **B,** a cluster at 9 – 14 Hz; 800 – 1000ms post stimulus onset. In both A and B, only electrodes for which the alpha effects were significantly different are marked.

II. Amount of alpha suppression modulated by social opinion ratings

The increase in alpha suppression observed for Interacted-with avatars suggests greater attentional resources were allocated to them than to the Non-interacted-with avatars. The second aim of this study is to determine whether the alpha modulation (i.e. attention allocation) varied as a function of how the avatars were rated on a certain trait.

The purpose of having the participants interact with three different avatars was to ensure as wide a range of ratings per trait as possible. Table 6.2 shows the spread of ratings for each of the three traits (perceived humanness, perceived strangeness, and quality of facial expression) where a higher rating represents a higher score in said trait (more human, more strange, higher quality of facial expression). Participants were asked, after interacting with each avatar, to rate said avatar on each of these traits *in relation to the other two avatars.* For this reason, participants were allowed to change their answers after having interacted with all three.

Table 6.2 Number of ratings in each Likert-Scale trait. Three traits were tested (perceived humanness, perceived strangeness and quality of facial expressions). Higher ratings represent a higher score in said trait (more human, more strange, higher quality of facial expression).

	Rating					
Trait	**1**	**2**	**3**	**4**	**5**	**6**
Perceived Humanness	1	18	22	37	10	2
Perceived Strangeness	7	34	25	21	3	0
Quality of Facial Expression	9	25	21	24	9	2

As predicted, we have a wide spread of ratings across most of the traits, with only the outer most ratings having 3 data points or less. Before using these data in any analysis, they were first trimmed to only include ratings with 7 or more data points. In other words, we removed rating 1 (N = 1) and 6 (N = 2) for perceived humanness, rating 5 (N = 3) and 6 (N = 0) for perceived strangeness, and rating 6 (N = 2) for quality of facial expression.

To determine if these ratings can modulate the alpha activity shown in Figure 6.3, we used the channels that showed significant activity (i.e., the ones indicated in Figure 6.3) as regions of interest. We extracted the average values for each participant for each avatar for the early and late alpha band frequency and time windows as specified above. The values for Phase 1 were then subtracted from the values for Phase 3 to control for any changes in activity related to the processing of photos. In the resulting dataset, the values for the Non-interacted-with avatars were then subtracted from the values for the Interacted-with avatars to control for the fact that the photos were viewed for a second time. The resulting values therefore capture any changes in alpha activity due to how participants process the Interacted-with avatars. These resulting values were then entered into a mixed model along with the trimmed behavioural ratings for each of the three traits in Table 6.2.

Only the model for the picture-induced alpha suppression (400 – 450ms; 11 – 14 Hz) showed a tendency for the Perceived Humanness rating to predict alpha activity ($\chi^2(1) = 3.55$, $p = .059$). Perceived Strangeness was in the best model, but showed no significant contribution ($\chi^2(1) = 0.20$, $p = .656$). Figure 6.4 shows the influence of Perceived Strangeness and Perceived Humanness on changes in alpha activity between Interacted-with and Non-interacted-with avatars.

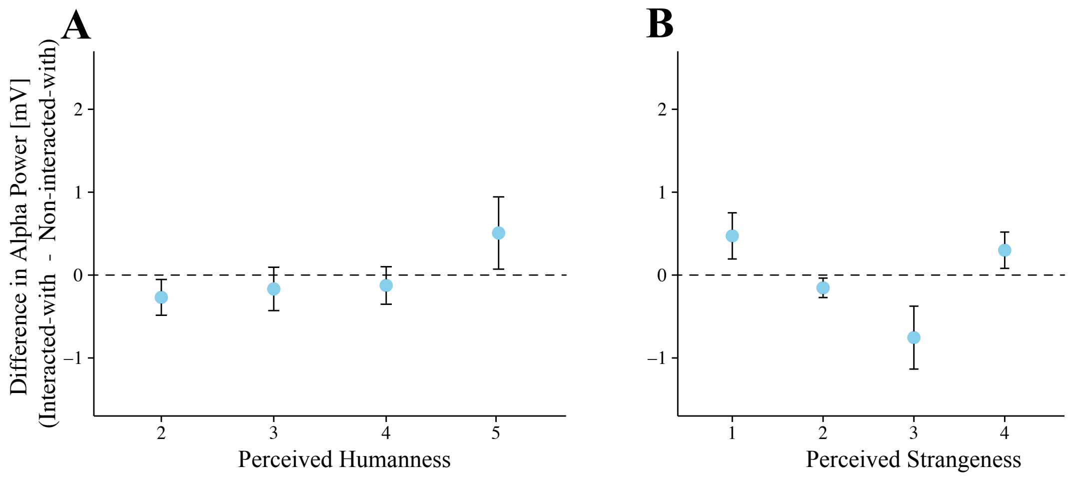

Figure 6.4 Difference in alpha activity between Perceived Humanness (A) and Perceived Strangeness (B) Ratings. Only ratings for perceived strangeness (modelled as a quadratic term) significantly predicted the difference in alpha power to the presentation of the faces of the Interacted-with avatar compared to the Non-interacted-with avatar. This suggests that the medially-strange rated avatar (rating of 3 out of 6) was allocated the most attention.

Due to the quadratic shape of the effect of Perceived Strangeness on alpha activity, we re-ran the mixed effects model with Perceived Strangeness modelled as a quadratic term. The results of this model are summarized in Table 6.3. The quadratic term significantly predicted the alpha suppression ($p = .010$). Previous studies have shown a modulation of language behaviour as a function of perceived strangeness. Our current result suggests that perception of this trait modulates attention allocation such that avatars which are perceived as being very strange or not very strange draw the least amount of attention, whereas those that are perceived as being medially-strange draw the most amount of attention. This could therefore explain the change in behaviour seen in previous studies.

Table 6.3 Summary of best linear mixed model for changes in alpha activity between Interacted-with and Non-interacted-with avatars.

Predictor	coefficient	*SE*	*t* value	*p*	
Intercept	-0.41	0.17	-2.49	.010	*
Perceived Strangeness (quad.)	0.01	0.04	2.56	.010	*
Perceived Humanness (linear)	-0.21	0.12	-1.78	.075	.
N = 84				. < .1 * < .05	

Discussion

In the current study we searched for neural evidence that attentional resources were differentially allocated when participants viewed faces of partners they had just interacted with and found that this is indeed the case. In addition, we examined if this attentional allocation differs as a function of the social opinion the participant has of their interaction partner and again found a correlation between evaluation ratings of the partners and modulations in alpha suppression. The term social opinion refers to the opinion one forms of their partner in social contexts (usually a positive or a negative opinion).

Our first aim was to determine whether attentional resources were differentially allocated when participants viewed faces of partners they had just interacted with. In order to test this, participants viewed pictures of digital partners before and after interacting and evaluating them while we recorded their EEG. The Non-interacted-with partners were partners that appeared in both parts of the EEG experiment but were not interacted with. To neurally gauge attentional allocation we examined the degree of alpha power suppression induced by the pictures of digital partners. We found that the degree of alpha suppression varied depending on interactions with the digital partners, and as a function of the ratings participants gave these partners. Specifically, we found that the Interacted-with avatars induced a greater amount of alpha suppression 400ms after picture onset compared to the Non-interacted-with avatar. Alpha suppression occurred over the occipital parietal cortices, and had a high peak frequency (11 – 14 Hz), which has been previously found to be associated with semantic processing demands (Kroll and Klimesch, 1992; Klimesch et al., 1994; Klimesch, 1997). Taken together, this would suggest that pictures of the Interacted-with avatars have been allocated more resources for processing compared to the Non-interacted-with avatars. The posterior location of this cluster is suggestive of a similar effect seen in analyzing emotional stimuli which indicates enhanced perceptual analysis of the input (De Cesarei & Codispoti, 2011) or enhanced motivation (Cui et al., 2013).

In addition, we found that the late post-stimulus rebound in alpha power was significantly less for the Interacted-with versus the Non-interacted-with avatars 800 – 1000ms after stimulus presentation. Previous studies have suggested that the late alpha increase relative to baseline from 500 – 1000ms may reflect the disengagement of attention after initial processing (Jensen et al., 2012). What is interesting, however, is that this disengagement was more pronounced for the Non-interacted-with avatars.

Our second aim in this study was to determine whether the degree of picture-induced alpha modulation (i.e. degree of attentional allocation) differed as a function of the social opinion the participant has of their interaction partner. All the avatars were designed to appear exactly the same, except for the facial expressions, which we have shown in Chapters 3 and 4 to be enough to vary the perception and social opinion of these avatars. However, as they were identical in appearance, we cannot rule out the possibility that the participant perceived each avatar not as individuals, but as the same avatar with varying moods. However, this does not change the interpretation of our results, as it is still the opinion one has of their partner, whether it is the same partner or different, that affects attentional allocation.

We found a modulation of alpha suppression as a function of the perceived strangeness rating in the form of a U-shaped curve for the early alpha effect (400 – 450ms). The avatars that were rated highest and lowest on the strangeness scale induced less alpha suppression in the Interacted-with compared to the Non-interacted-with avatars, whereas the middle rating elicited less alpha suppression for the Non-interacted-with compared to the Interacted-with avatars. This suggests that at the extremes, perception of the Interacted-with avatars requires less attentional resources compared to perception of the Non-interacted-with avatars.

The avatars we used in this study have been used in Chapter 4 to elicit a difference in syntactic priming as a function of the strangeness rating. This effect took the form of an inverted U-shaped curve such that the maximal syntactic priming effect occurred for the middle ratings of perceived strangeness, with a decrease on either side of this middle rating. Together with the EEG data from the current study it is clear that maximal syntactic priming occurs when the participant is allocating more attention to their conversation partner. As syntactic priming is theorized to be supported by implicit learning (Chang et al., 2006; Jaeger and Snider, 2013), this would suggest that more attention allocated to the partner results in more robust representations of what is being said, to allow for the participant to (implicitly) learn these utterances well enough to use them in the future.

Our results also shed light on theories about why performance differs when participants do the same task with a partner compared to when they do it without a partner. We do not find direct support for any theory postulating how attention can influence behaviour when interacting with a partner. The closest is the feedback-loop model, which suggests that interacting with a person causes the participant to focus more on the self (Duval and Wicklund, 1973; Wicklund, 1975; for a recent review see Strauss, 2002). Our results, however, suggest the opposite. The decreased alpha power for the Interacted-with versus the Non-interacted-with avatars should have resulted in less syntactic priming if participant assigned more attention to the self. Syntactic priming only occurs when the information received from the partner is properly encoded; hence if less attention is allocated to the partner and more to the self, this information is encoded less and should therefore result in a lower priming magnitude. As this is not what we found, we suggest that the decreased alpha power, and therefore the increased attention, assigned for the Interacted-with avatars actually reflects more attention to the partner, not the self.

The difference in attentional resource allocation between Interacted-with and Non-interacted-with avatars suggests that having an association with a partner (in this study through joint action) results in more sensory processing when viewing pictures of this partner at a later point in time. This additional processing could reflect memory retrieval as the participants have interacted with them previously. It could also be related to having a social opinion compared to having no social opinion. As forming a social opinion does involve some aspects of emotion, previous studies have shown that items that elicit emotions require more processing resources compared to neutral items (Balconi et al., 2009; de Cesarei and Codispoti, 2011; Cui et al., 2013). The first aim of our study was to determine how attentional resources are allocated when interacting with a person. Our second aim was to determine whether this allocation of attention differs as a function of rating, and, again, we show that it does. This therefore clarifies why participant behaviour is different depending on whether they positively or negatively view their partner.

Social opinion is most likely supported by a combination of perceptual, cognitive, and motor operations and it is therefore not easy to disentangle what underlying features influence behaviour. Our study was only a first step in determining the neural mechanisms related to the influence of social opinion on behaviour, providing evidence for the role of attention. However, future studies will need to future tease apart what other features make up social opinion.

7

General Discussion

The aim of this thesis was to explore how different processes support and influence sentence processing by using a syntactic priming task. We measured syntactic priming tendency by using a picture description task, a methodology commonly used in the linguistic priming literature (Bock, 1986a; Bock & Loebell, 1990; Branigan et al., 2000; Ferreira et al., 2008; Segaert et al., 2011; Weatherholtz et al., 2014; inter alia) which hence provides a rich literature to build upon. Most studies have focused on the act of priming itself by investigating the characteristics of cumulativity and surprisal (Jaeger and Snider, 2008; Segaert et al., 2011), the rate of decay (Branigan et al., 1999; Wheeldon and Smith, 2003), the role of the lexical boost (Hartsuiker et al., 2008), and so forth. However, sentence comprehension and production most commonly occur in a dialogue setting, where extrinsic factors such as the opinion one has of their partner have been suggested to modulate how well we process syntactic structure (Pickering and Garrod, 2004, 2013; Balcetis and Dale, 2005; Weatherholtz et al., 2014). Therefore, this thesis aims to investigate how such extrinsic factors could influence sentence processing, and touches upon processes such as memory and attention during the course of this investigation.

The Role of Memory in Syntactic Priming

Current theories of syntactic priming have suggested that the behavioural result of repeating grammatical structures across utterances is supported by both the explicit and the implicit memory system (Bock and Griffin, 2000; Chang et al., 2006; Ferreira and Bock, 2006). However, testing this theory in healthy individuals is difficult as it is hard to separate contributions of either memory system. As patients with amnesia are impaired in their explicit memory system, they represent an ideal patient group to investigate the influences of either explicit or implicit memory on a certain behaviour.

In terms of syntactic priming, we were not the first to use patients to study the influence of memory on priming ability. Ferreira et al. (2008) had previously published a study in which they compared four patients with amnesia to four age- and education-matched controls on a syntactic priming task. In Chapter 2 we aimed to replicate the effect in a more homologous patient population with a larger sample size. Overall we were able to show that our patient group does prime, supporting the role of implicit memory in syntactic priming.

Unexpectedly, we discovered that there may be an age-related decline in syntactic processing as our control group did not prime. We originally recorded over 40 participants, and still we observed the same result: an absence of syntactic priming in healthy elderly participants. An age-related decline in syntactic processing has not been thoroughly explored, although age-related decline in language behaviour in general has been observed before (for review see Burke & Mackay, 1997).

The Role of Social Influences on Syntactic Priming

The rest of this thesis is dedicated to investigating how the opinion a participant has of their partner, referred to in this thesis as social opinion, could influence that participant's behaviour, such as their syntactic processing ability. As discussed in Chapters 3 and 4, there already exist a couple of studies, namely by Balcetis and Dale (2005) and Weatherholtz and colleagues (2014), that have investigated this for syntactic priming. However, the results contradicted each other, with Balcetis and Dale showing an increase in priming magnitude as the partner was rated more favourably by the participant, whereas Weatherholtz and colleagues discovered a decrease in priming magnitude (for the same syntactic structure) with increased favourability ratings. With too many differences between the studies to isolate where these contradictory results may originate from, we conducted our own study in the hopes of clarifying the effect of social opinion on syntactic priming magnitude, and hence the effect of social opinion on sentence processing.

We discovered that the effect is not necessarily linear, as the Balcetis and Dale, and Weatherholtz studies observed. Instead, our results illustrated an inverted U-shaped curve as a function of partner rating. As our "partner" was a digital confederate ("avatar"), we found this effect in the perceived strangeness rating. For this rating, participants filled in a 6-point Likert scale question asking *how strange do you find this avatar in relation to the other avatars*. For this reason, we did not z-score the answers (which is common practice with questionnaire data) as the scores were already relative. We also included the questionnaire used in the Weatherholtz study to evaluate the avatar partner. In the Weatherholtz study, the interaction with priming magnitude was mainly found on the *Interpersonal Similarity* factor, which presented itself after factor analysis. We were able to observe the same factor, with the same questions loading similarly compared to the Weatherholtz study, however, we did not find a direct effect of this loading on priming magnitude. We believe this might be due to the computer-ness of our avatar. As participants knew that they were conversing with a computer, answering a questionnaire designed to evaluate human partners may have been a bit strange, and hence no direct link to behaviour was found. We did see an indirect link, as *Interpersonal Similarity* was correlated with the rating of perceived strangeness.

We therefore confirm that syntactic priming, and hence sentence processing, is modulated by the opinion one has of their partner. We also bring together the contradictory results seen in the human literature, as our inverted U-shaped curve suggests that previous studies may have been measuring from different parts of the curve, and hence measured either a positive or a negative influence.

However, Chapter 3 and 4 do not explain *why* social opinion modulates priming magnitude. The opinion one has of their partner is a single behavioural (and abstract) measurement that is most likely supported by a mixture of underlying processes in the brain. Hence for the rest of the thesis we aimed to determine what these underlying mechanisms might be.

The Role of Attention on Syntactic Priming

One likely component to influence the forming of an opinion is attention. Having a person be more or less distracting, perhaps through body language, speech habits, etc. may implicitly influence the evaluation of that person. However, the first step is to determine if syntactic priming magnitude can be modulated by attention. Therefore in Chapter 5 we used a dual-task paradigm to determine if syntactic processing is vulnerable to modulations in attentional resources. By using a domain-general secondary task such as the motion-object tracking (MOT) task, we were able to determine whether non-language-specific distractions, such as keeping track of a set of moving balls, could indeed influence syntactic priming. In Chapter 4, we manipulated the opinion participants formed of the avatar by manipulating the facial expressions the avatar presented. Therefore, to be able to draw parallels between the dual-task in Chapter 5 and the avatar facial expressions in Chapter 4, we wanted only visual cues to modulate attention.

Our dual-task paradigm indeed showed that the MOT task can influence the magnitude of the syntactic priming effect, providing evidence that attention does influence syntactic priming. Our results suggest that our attention demanding task affected the encoding of the syntactic information, not the retrieval, as we only observed a modulation in the magnitude of the priming effect when the MOT task was done concurrently with the presentation of the prime picture (when participants listened to the picture description and hence when they encode the syntactic information; Encoding Phase). We observed no effect on syntactic priming magnitude when the MOT task was done concurrently with the presentation of the target picture (when the participants describe the picture and hence retrieve syntactic information; Retrieval Phase).

Looking at the accuracy of the MOT task responses shows a significant decrease in the ability to identify the ball to-be-tracked when doing the MOT task concurrently with the language task compared to doing the task alone. As we did not see a difference in syntactic priming magnitude in the dual-task compared to the single-task condition, the data suggests that the language task was given the priority of the resources. However, this is contradictory to what has been shown in driving studies in which the participant drives while conversing (Becic et al., 2010). In these studies, the data suggests that driving gets the priority of the resources. Although our data contradicts these claims, it is possible that context (i.e., driving is more dangerous than following balls on a screen) can influence the priority of the task. Therefore, even though we do not see an effect of attention modulation on syntactic priming magnitude in the Retrieval Phase in our study, if the secondary task was more important (like driving) perhaps we would see different results.

Our study provided evidence that attention can influence the magnitude of the syntactic priming effect. However, to bring this back to the theory that attention could be a process influencing social opinion, our results in Chapter 5 provide only an indirect connection to social opinion. Attention is a component which modulates syntactic processing, but there is no evidence to suggest that it also influences social opinion. Hence for the final chapter of this thesis, we aimed to provide evidence that modulations in attention do result in modulations in social opinion.

In this final chapter (Chapter 6) we used electroencephalography (EEG) to determine whether attentional allocation is modulated depending on the ratings the participants gave the avatars. In Chapter 4 we only observed an effect of the perceived strangeness rating on syntactic priming magnitude. Participants answered the question *how strange is this avatar compared to the other two avatars?* for each of the three avatars they interacted with. When correlated to the magnitude of the syntactic priming effect, we observed a quadratic curve such that the middle rated avatar elicited the highest priming magnitude. In our final chapter we also observed a modulation in the allocation of attentional resources as a function of the perceived strangeness rating such that more attention was allocated to the middle rated avatar. As both observations were quadratic curves, this strongly suggests that we managed to observe the same effect in EEG as we did behaviourally in Chapter 4 (Figure 7.1). We thus have been able to not only observe how social opinion can influence syntactic priming, we were also able to provide neural evidence to support this observation.

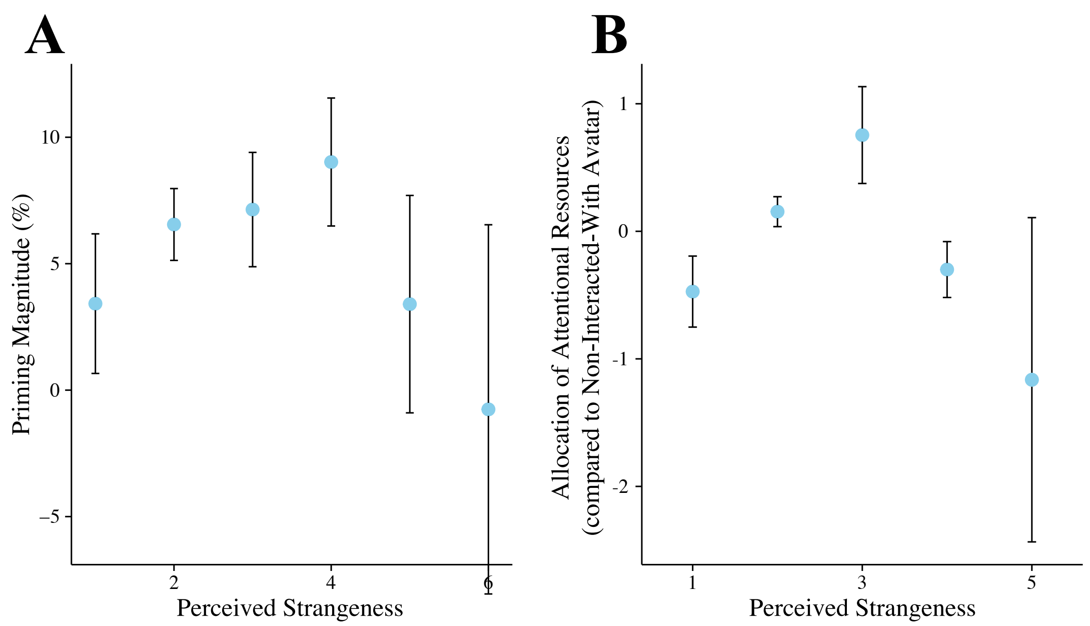

Figure 7.1 The quadratic curves seen in Chapter 4 (A) and Chapter 6 (B; y axis inverted relative to original) show similar trends along the Perceived Strangeness rating. This suggests that the allocation of attentional resources is the reason priming magnitude is modulated along Perceived Strangeness.

The Role of Implicit Memory – Closing the Loop

The changes in alpha oscillations not only show that attention is modulated as a function of social opinion, but it also illustrates how this affects priming magnitude. Figure 7.1B shows the difference in the allocation of attentional resources between Interacted-with and Non-interacted-with avatars. A value of 0 on this scale therefore suggests that there is no difference in the allocation of attention when viewing a partner you have previously interacted with versus viewing a partner you have not previously interacted with. This value can therefore be viewed as the baseline attentional allocation amount. This value roughly corresponds to a priming magnitude (in our study) of 6% (Figure 7.1A). Decreasing the amount of attention allocated results in a decrease in priming magnitude, whereas an increase in attentional allocation results in an increase in priming magnitude. Our results therefore support that attentional allocation can not only decrease priming magnitude in cases such as distraction, but it can also *increase* priming magnitude by drawing more attention towards the interaction partner.

We additionally suggest that the different processes that influence syntactic priming (social opinion, attention allocation, etc.) are influences on the workings of implicit learning that supports syntactic priming. Increased attention to the interaction partner ensures increased encoding of syntactic information received from that partner. We show evidence for this in both our dual-task study in Chapter 5, as well as the EEG study in Chapter 6. Therefore, it is not that implicit memory is one system that could influence syntactic priming magnitude, it is *the* system that influences syntactic priming magnitude (Figure 7.2).

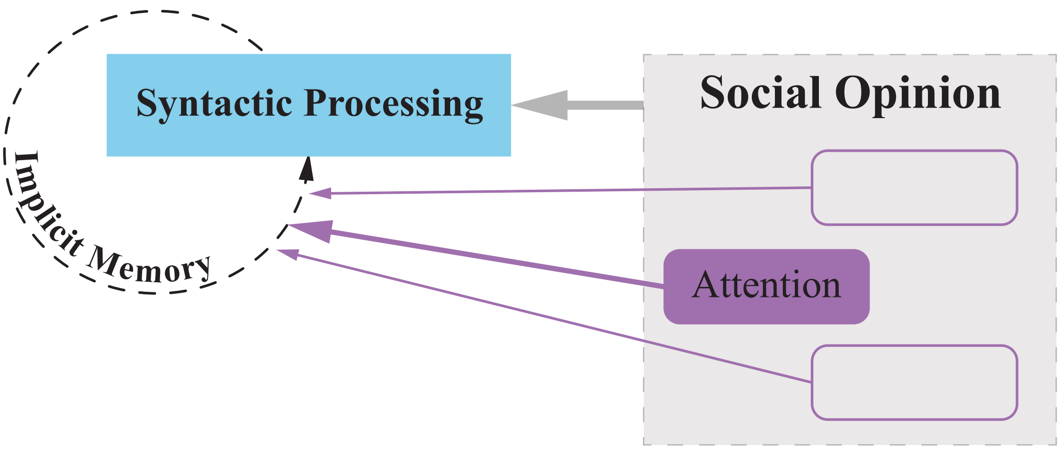

Figure 7.2 Previous studies suggested that syntactic processing is influenced by the opinion one has of their partner (social opinion). After concluding this thesis, we propose that social opinion is made up of different components (one of them being attention) that influence syntactic processing not directly, but via implicit learning.

Remaining Mysteries

We have provided evidence indicating that modulations in attention are correlated with modulations in syntactic priming. However, one question that prominently remains is what causes this modulation in attention to begin with? Previous studies have shown a modulation in syntactic priming as a function of "interpersonal similarity." It is possible that partners that are perceived as being either very similar or not similar at all, for some reason are less interesting and hence less attention is assigned to them. This argument seems intuitive for partners that are perceived as being very similar to the participant; this similarity to oneself could make it unnecessary to allocate a lot of attention to the partner's characteristics. The opposite, however, is harder to explain, although it could simply be that if the discrepancy between oneself and one's partner is too large, perhaps it is not worth assigning the additional attentional resources to further evaluate the partner at the risk of being worse at the task. This idea is supported by Chapter 5, where we show that when one is minimally distracted, syntactic priming magnitude is actually higher compared to being not taxed at all, or being overly taxed. In Chapter 5 we explain this "boost" as an increase in perceptual processes that also causes enhanced encoding of syntactic information. It is possible that in interaction situations, the "boost" is also caused when the partner is a little distracting (so neither too similar nor too dissimilar), which results in enhanced encoding of syntactic information. Of course, further studies are needed to further develop this theory.

The connection between attention and syntactic priming has only been explored in one study using a dual-task methodology. However, there are multiple other methodologies that can be used to verify this connection. One way that has already been investigated is the use of patients with autism spectrum disorder (ASD). ASD patients do not attend to their interaction partner to the same extent as a neurologically typical individual as usually ASD patients focus more on themselves and how they conduct the task. Recent studies have shown that ASD patients do prime significantly less compared to neurologically typical controls, at least in the realm of behavioural priming (Vivanti and Hamilton, 2014; Forbes et al., 2016).

We are not suggesting that attention is the only component that could influence how social opinion influences sentence processing. As sentence processing is supported by statistical learning, it could be that the predictability of the interaction partner also influences how fast and/or to what extent the participant primes to their interaction partner. Previous studies have already shown that by reducing the surprisal effect, and therefore influencing the predictability of the grammatical structures, the magnitude of the passive priming effect decreases (Segaert et al., 2011). Therefore, reducing the surprisal effect by making the interaction partner more predictable might also influence how much the participant primes to that interaction partner.

Finally, we suggested that syntactic priming may be prone to age-related decline. This line of research is only just starting to be investigated (Sung, 2015, 2016) and hence more work is necessary to determine if it is truly due to age-related decline in syntactic processing and/or implicit memory. Indeed, as theories of syntactic priming have suggested that its function is to aid the efficiency of language by re-using previously processed utterances, it would more logical that if there was an age-related decline in sentence processing, it should result in an increase in priming, rather than a decrease. Therefore, age-related declines may more likely be linked to age-related deficits in implicit memory and statistical learning, which goes against studies suggesting that implicit memory does not decline with age (Burke and Mackay, 1997; Schugens et al., 2007). Additionally, it could be that the task we are using to assess syntactic priming requires additional processes beyond just sentence processing and implicit memory (for example, working memory) and that these additional processes show an age-related decline.

To summarize, in this thesis I explored whether social opinion does indeed modulate syntactic processing, and how. Looking at all the findings in this thesis, it seems most likely that implicit memory plays a major role in influencing syntactic processing and that any apparent other influences, such as social opinion, are rather influences on implicit learning than influences on sentence processing mechanisms.

R

References

Allen R, Mcgeorge P, Pearson D, Milne AB (2004) Attention and expertise in multiple target tracking. Appl Cogn Psychol 18:337–347.

Alsius A, Navarra J, Campbell R, Soto-Faraco S (2005) Audiovisual integration of speech falters under high attention demands. Curr Biol 15:839–843.

Altmann GTM, Kamide Y (1999) Incremental interpretation at verbs: Restricting the domain of subsequent reference. Cognition 73:247–264.

Andonova E (2013) Gender and Sex: The Experiential Basis of Grammar. In: Proceedings of the 35th Annual Conference of the Cognitive Science Society, pp 1738–1743.

Angwin AJ, Chenery HJ, Copland DA, Arnott WL, Murdoch BE, Silburn PA (2004) Dopamine and semantic activation: an investigation of masked direct and indirect priming. J Int Neuropsychol Soc 10:15–25.

Bäckman L, Nyberg L, Lindenberger U, Li SC, Farde L (2006) The correlative triad among aging, dopamine, and cognition: current status and future prospects. Neurosci Biobehav Rev:791–807.

Balcetis EE, Dale R (2005) An Exploration of Social Modulation of Syntactic Priming. In: Proceedings of the 27th Annual Meeting of the Cognitive Science Society (Bara BG, Barsalou L, Bucciarelli M, eds), pp 184–189. Stresa, Italy.

Balconi M, Falbo L, Brambilla E (2009) BIS/BAS responses to emotional cues: Self report, autonomic measure and alpha band modulation. Pers Individ Dif 47:858–863.

Bargh JA (2012) Priming Effects Replicate Just Fine, Thanks. Psychol Today.

Bargh JA, Chen M, Burrows L (1996) Automaticity of social behavior: direct effects of trait construct and stereotype-activation on action. J Pers Soc Psychol 71:230–244.

Baron RS (1986) Distraction-conflict theory: Progress and problems. Adv Exp Soc Psychol 19:1–40.

Barr DJ, Levy R, Scheepers C, Tily HJ (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. J Mem Lang 68:255–278

Bates D, Maechler M, Bolker B (2012) lme4: Linear mixed-effects models using S4 classes (2011). R package version 0.999375-42.

Becic E, Dell GS, Bock JK, Garnsey SM, Kubose T, Kramer AF (2010) Driving impairs talking. Psychon Bull Rev 17:15–21.

Beckner C, Racz P, Hay J, Brandstetter J, Bartneck C (2015) Participants Conform to Humans But Not to Humanoid Robots in an English Past Tense Formation Task. J Lang Soc Psychol:1–22.

Bee N, André E, Tober S (2009) Breaking the Ice in Human-Agent Communication : Eye-Gaze Based Initiation of Contact with an Embodied Conversational Agent. In: Intelligent Virtual Agents, pp 229–242. Springer Berlin Heidelberg.

Bergmann K, Branigan HP, Kopp S (2015) Exploring the Alignment Space - Lexical and Gestural Alignment with Real and Virtual Humans. Front ICT 2:1–11.

Berkum JJA Van, Brink D Van Den, Tesink CMJY, Kos M, Hagoort P (2008) The Neural Integration of Speaker and Message. J Cogn Neurosci 20:580–591.

Bernolet S, Collina S, Hartsuiker RJ (2016) The persistence of syntactic priming revisited. J Mem Lang.

Blascovich J, Loomis J, Beall AC, Swinth KR, Hoyt CL, Bailenson JN (2002) Immersive Virtual Environment Technology as a Methodological Tool for Social Psychology. 13:103–124.

Bock JK (1986a) Syntactic Persistence in Language Production. Cogn Psychol 18:355–387.

Bock JK (1986b) Meaning, sound, and syntax: Lexical priming in sentence production. J Exp Psychol Learn Mem Cogn 12:575–586.

Bock JK, Dell GS, Garnsey SM, Kramer AF, Kubose T. (2007) Car talk, car listen. In: Automaticity and Control in Language Processing, pp 21–42. Taylor & Francis.

Bock JK, Griffin ZM (2000) The Persistence of Structural Priming: Transient Activation or Implicit Learning? J Exp Psychol Gen 129:177–192.

Bock JK, Loebell H, Morey R (1992) From conceptual roles to structural relations: Bridging the syntactic cleft. Psychol Rev 99:150–171.

Bond CF, Titus LJ (1983) Social Facilitation: A Meta-Analysis of 241 Studies. Psychol Bull 94:265–292.

Bourke PA (1996) A General Factor Involved in Dual task Performance Decrement. Q J Exp Psychol Sect A 49:525–545.

Bowers B (2012) The hot and cold of priming: Psychologists are divided on whether unnoticed cues can influence behavior. Sci News 181:26–29.

Branigan HP, Pickering MJ, Cleland AA (1999) Syntactic priming in written production: evidence for rapid decay. Psychon Bull Rev 6:635–640.

Branigan HP, Pickering MJ, Pearson J, McLean JF, Nass CI (2003) Syntactic Alignment Between Computers and People: The Role of Belief about Mental States. In: Proceedings of the 25th Annual Conference of the Cognitive Science Society (Alterman R, Kirsh D, eds), pp 186–191. Boston, MA.

Branigan HP, Pickering MJ, Stewart A, McLean JF (2000) Syntactic priming in spoken production: linguistic and temporal interference. Mem Cognit 28:1297–1302.

Bühler K (1934) Sprachtheorie [Theory of Language].

Burke DM, Mackay DG (1997) Memory, language, and ageing. Philos Trans R Soc Lond B Biol Sci 352:1845–1856.

Casasanto LS, Jasmin K, Casasanto D (2010) Virtually accommodating : Speech rate accommodation to a virtual interlocutor. In: Proceedings of the 32nd Annual Conference of the Cognitive Science Society. (Camtrabone R, Ohlsson S, eds), pp 127–132. Portland, OR.

Cermak LS, Verfaellie M, Milberg W, Letourneau L, Blackford S (1991) A further analysis of perceptual identification priming in alcoholic Korsakoff patients. Neuropsychologia 29:725–736.

Chang F (2002) Symbolically speaking : a connectionist model of sentence production. Cogn Sci 26:609–651.

Chang F, Dell GS, Bock JK (2006) Becoming syntactic. Psychol Rev 113:234–272.

Chang F, Dell GS, Bock JK, Griffin ZM (2000) Structural Priming as Implicit Learning: A Comparison of Models of Sentence Production. J Psycholinguist Res 29:217–229.

Charny IW (1966) Integrated individual and family psychotherapy. Fam Process 5:179–198.

Chartrand TL, Bargh JA (1999) The Chameleon Effect: The Perception-Behavior Link and Social Interaction. J Pers Soc Psychol 76:893–910.

Conway C, Christiansen M (2001) Sequential learning in non-human primates. Trends Cogn Sci 12:539–546.

Coyle JM, Kaschak MP (2012) Female fertility affects men's linguistic choices. PLoS One 7.

Cui Y, Versace F, Engelmann JM, Minnix JA, Robinson JD, Lam CY, Karamhage M, Brown VL, Wetter DW, Dani JA, Kosten TR, Cinciripini PM (2013) Alpha Oscillations in Response to Affective and Cigarette-Related Stimuli in Smokers. Nicotine Tob Res 15:917–924.

Dan-glauser ES, Scherer KR (2011) The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance. Behav Res Methods 43:468–477.

de Cesarei A, Codispoti M (2011) Affective modulation of the LPP and α-ERD during picture viewing. Psychophysiology 48:1397–1404.

De Renzi E (1989) Apraxia. In: Handbook of neuropsychology, 2nd ed. (Boller F, Grafman J, eds), pp 245–263. New York: Elsevier Science.

Dehaene S, Meyniel F, Wacongne C, Wang L, Pallier C (2015) The Neural Representation of Sequences: From Transition Probabilities to Algebraic Patterns and Linguistic Trees. Neuron 88:2–19 Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9–21.

Deubel H, Schneider WX (1996) Saccade target selection and object recognition: Evidence for a common attentional mechanism. Vision Res 36:1827–1837.

Dijksterhuis AP, Bargh JA (2001) The perception-behavior expressway: Automatic effects of social perception on social behaviour. In: Advances in experimental social psychology (Zanna MP, ed), pp 1–40. San Diego: Academic Press.

Dijksterhuis AP, Spears R, Postmes T, Stapel D, Koomen W, Knippenberg A Van, Scheepers D (1998) Seeing one thing and doing another: Contrast effects in automatic behavior. J Pers Soc Psychol 75:862–871.

DSM-5 American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders. Arlington: American Psychiatric Publishing.

Duval S, Wicklund RA (1973) Effects of objective self-awareness on attribution of causality. J Exp Soc Psychol 9:17–31.

Eichenbaum H, Cohen NJ (2001) From conditioning to conscious recollection: memory systems of the brain. New York: Oxford University Press.

Ferreira VS (2003) The persistence of optional complementizer production: Why saying “that” is not saying “that” at all. J Mem Lang 48:379–398.

Ferreira VS, Bock JK (2006) The functions of structural priming. Lang Cogn Process 21:1011–1029.

Ferreira VS, Bock JK, Wilson MP, Cohen NJ (2008) Memory for syntax despite amnesia. Psychol Sci 19:940–946.

Fincher-Kiefer R (2001) Perceptual components of situation models. Mem Cognit 29:336–343.

Fincher-Kiefer R, D’Agostino PR (2004) The role of visuospatial resources in generating predictive and bridging inferences. Discourse Process 37:205–224.

Finlayson IR, Corley M (2012) Disfluency in dialogue: an intentional signal from the speaker? Psychon Bull Rev 19:921–928

Fogg BJ, Nass CI (1997) How users reciprocate to computers: an experiment that demonstrates behavior change. In: CHI’97 extended abstracts on Human factors in computing systems. ACM.

Forbes P, Pan X, Hamilton A (2016) Reduced mimicry to virtual reality avatars in autism. In: The 2nd Virtual Social Interaction Workshop.

Fougnie D, Marois R (2006) Distinct capacity limits for attention and working memory: Evidence from attentive tracking and visual working memory paradigms. Psychol Sci 17:526–534

Fox J, Weisberg S (2011) Car: companion to applied regression.

Foxe JJ, Simpson G V, Ahlfors SP (1998) Parieto-occipital- 1 0Hz activity reflects anticipatory state of visual attention mechanisms. Neuroreport 9:3929–3933.

Foxe JJ, Snyder AC (2011) The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. Front Psychol 2:1–13.

Friederici AD (2002) Towards a neural basis of auditory sentence processing. 6:78–84.

Giles H, Coupland J, Coupland N (1991) Contexts of accommodation: Developments in applied sociolinguistics. Cambridge University Press.

Giles H, Henwood K, Coupland N, Harriman J, Coupland J (1992) Language attitudes and cognitive mediation. Hum Commun Res 18:500–527.

Giles H, Powesland PF (1975) Speech styles and social evaluation. Academic Press.

Gong L (2008) How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. Comput Human Behav 24:1494–1509.

Gries ST (2005) Syntactic priming: A corpus-based approach. J Psycholinguist Res 34:365–399.

Hagoort P (2003) How the brain solves the binding problem for language: a neurocomputational model of syntactic processing. Neuroimage:S18–S20.

Hagoort P (2005) On Broca, brain, and binding: a new framework. Trends Cogn Sci 9:416–423

Hagoort P (2013) MUC (Memory, Unification, Control) and beyond. Front Psychol 4:416 A

Hagoort P (2016) MUC (Memory, Unification, Control): A Model on the Neurobiology of Language Beyond Single Word Processing. In: Neurobiology of language (Hickok G, Small S, eds), pp 339–347. Amsterdam: Elsevier.

Harley TA (1984) A critique of top-down independent levels models of speech production: Evidence from non-plan-internal speech errors. Cogn Sci 8:191–219.

Hartsuiker RJ, Bernolet S, Schoonbaert S, Speybroeck S, Vanderelst D (2008) Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. J Mem Lang 58:214–238

Hartsuiker RJ, Kolk HH (1998a) Syntactic facilitation in agrammatic sentence production. Brain Lang 62:221–254

Hartsuiker RJ, Kolk HHJ (1998b) Syntactic Persistence in Dutch. Lang Speech 41:143–184.

Herman CP (2015) The social facilitation of eating. A review. Appetite 86:61–73 Available at: http://dx.doi.org/10.1016/j.appet.2014.09.016.

Hess U, Philippot P, Blairy S (1999) Mimicry: Facts and fiction (Philippot P, Feldman RS, Coats EJ, eds). New York, NY, USA: Cambridge University Press.

Heyselaar E, Hagoort P, Segaert K (2015) In dialogue with an avatar, language behavior is identical to dialogue with a human partner. Behav Res Methods:1–15.

Hodges JR, Patterson K (1997) Semantic memory disorders. Trends Cogn Sci 2:68–72.

Howard D V, Heisey JG, Shaw RJ (1986) Aging and the Priming of Newly Learned Associations. 22:78–85.

Jackendoff R (2002) Foundations of language: brain, meaning, grammar, evolution. Oxford, UK: Oxford University Press.

Jaeger TF (2009) Random effect: Should I stay or should I go? Post to HLP/ Jaeger lab blog:http://hlplab.wordpress.com/2009/05/14/random-effe.

Jaeger TF, Snider NE (2008) Implicit learning and syntactic persistence : Surprisal and Cumulativity. Univ Rochester Work Pap Lang Sci 3:26–44.

Jaeger TF, Snider NE (2013) Alignment as a consequence of expectation adaptation: syntactic priming is affected by the prime's prediction error given both prior and recent experience. Cognition 127:57–83

Jensen O, Bonnefond M, VanRullen R (2012) An oscillatory mechanism for prioritizing salient unattended stimuli. Trends Cogn Sci 16:200–205 Available at: http://dx.doi.org/10.1016/j.tics.2012.03.002.

Johnson PCD, Barry SJE, Ferguson HM, Müller P (2015) Power analysis for generalized linear mixed models in ecology and evolution. Methods Ecol Evol 6:133–142.

Jollife IT (1972) Discarding variables in a principal component analysis, I: Artificial data. Appl Stat 21.

Jollife IT (1986) Principal components analysis. New York: Springer.

Kaschak MP, Kutta TJ, Jones JL (2011a) Structural priming as implicit learning: cumulative priming effects and individual differences. Psychon Bull Rev 18:1133–1139

Kaschak MP, Kutta TJ, Schatschneider C (2011b) Long-term cumulative structural priming persists for (at least) one week. Mem Cognit 39:381–388

Kelly SP, Lalor EC, Reilly RB, Foxe JJ (2006) Increases in alpha oscillatory power reflect an active retinotopic mechanism for distracter suppresion during sustained visuospatial attention. J Neurophysiol 95:3844–3851.

Kessels RPC, Remmerswaal M, Wilson B a (2011) Assessment of nondeclarative learning in severe Alzheimer dementia: the Implicit Memory Test (IMT). Alzheimer Dis Assoc Disord 25:179–183.

Kidd E (2012) Implicit statistical learning is directly associated with the acquisition of syntax. Dev Psychol 48:171–184

Kinchla RA (1992) Attention. Annu Rev Psychol 43:711–742.

Klimesch W (1997) EEG-alpha rhythms and memory processes. Int J Psychophysiol 26:319–340.

Klimesch W, Sauseng P, Hanslmayr S (2007) EEG alpha oscillations: the inhibition-timing hypothesis. Brain Res Rev 53:63–88.

Klimesch W, Schimnke H, Schwaiger J (1994) Episodic and semantic memory: an analysis in the EEG-theta and alpha band. Clin Neurophysiol 91:428–441.

Knopman D, Nissen MJ (1994) Procedural learning is impaired in Huntington's disease: evidence from the serial reaction time task. Neuropsychologia:245–254.

Kopelman MD (2002) Disorders of memory. Brain:2152–2190.

Kopelman MD (2014) What does a comparison of the alcoholic Korsakoff syndrome and thalamic infarction tell us about thalamic amnesia? Neurosci Biobehav Rev 54:46–56

Korvorst M, Roelofs A, Levelt WJM (2006) Incrementality in naming and reading complex numerals: Evidence from eyetracking. Q J Exp Psychol 59:263–311.

Koulouri T, Lauria S, Macredie RD (2014) Do (and Say) as I Say: Linguistic Adaptation in Human–Computer Dialogs. Human–Computer Interact 24:1–37

Kroll NE, Klimesch W (1992) Semantic memory: complexity or connectivity? Mem Cognit 20:192–210.

Krupa AK (2009) The competitive nature of declarative and nondeclarative memory systems: Converging evidence from animal and human brain studies. UCLA Undergrad Sci J 22:39–46.

Kubose TT, Bock JK, Dell GS, Garnsey SM, Kramer AF, Mayhugh J (2006) The effects of speech production and speech comprehension on simulated driving performance. Appl Cogn Psychol 20:43–63.

Kuhlen AK, Brennan SE (2013) Language in dialogue: when confederates might be hazardous to your data. Psychon Bull Rev 20:54–72.

Lang PJ, Bradley MM, Cuthbert BN (1997) Motivated attention: Affect, activation, and action. Atten orienting Sens Motiv Process:97–135.

Langlais PJ, Mair RG, Anderson CD, Mcentee WJ (1987) Monoamines and metabolites in cortex and subcortical structures: normal and regional distribution and the effects of thiamine deficiency in the rat. Brain Res 1:140–149.

Lavie N, Tsal Y (1994) Perceptual load as a major determinant of. Percept Psychophys 56:183–197.

Lee E-J (2010) The more humanlike, the better? How speech type and users' cognitive style affect social responses to computers. Comput Human Behav 26:665–672.

Levelt WJM, Kelter S (1982) Surface Form and Memory in Question Answering. Cogn Psychol 14:78–106.

Looser CE, Wheatley T (2010) The tipping point of animacy: How, when, and where we perceive life in a face. Psychol Sci 21:1854–1862.

Lott AJ, Lott BE (1961) Group cohesiveness, communication level, and conformity. J Abnorm Soc Psychol 62:408.

Makovski T, Swallow KM, Jiang Y V. (2011) Attending to unrelated targets boosts short-term memory for color arrays. Neuropsychologia 49:1498–1505

Mandell AR, Smith MA, Martini MC, Shaw TH, Wiese E (2015) Does the Presence of Social Agents Improve Cognitive Performance on a Vigilance Task? In: International Conference on Social Robotics. Springer International Publishing, pp 421–430.

Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods 164:177–190.

Maurer RE, Tindall JH (1983) Effect of postural congruence on client's perception of counselor empathy. J Couns Psychol 30:158–163.

Mazaheri A, Jensen O (2010) Rhythmic pulsing: linking ongoing brain activity with evoked responses. Front Hum Neurosci 4.

Mazaheri A, van Schouwenburg MR, Dimitrijevic A, Denys D, Cools R, Jensen O (2014) Region-specific modulations in oscillatory alpha activity serve to facilitate processing in the visual and auditory modalities. Neuroimage 87:356–362.

McGurk H, MacDonald J (1976) Hearing lips and seeing voices. Nature 265:746–748.

Melo C de, Gratch J, Carnevale P (2014) Humans vs. Computers: impact of emotion expressions on people's decision making. :1–11.

Menenti L, Gierhan SME, Segaert K, Hagoort P (2011) Shared language: overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. Psychol Sci 22:1173–1182.

Mesulam M-M, Thompson CK, Weintraub S, Rogalski EJ (2015) The Wernicke conundrum and the anatomy of language comprehension in primary progressive aphasia. Brain 138:2423–2437

Meyer AS, Sleiderink AM, Levelt WJM (1998) Viewing and naming objects: Eye movements during noun phrase production. Cognition 66:B25–B33.

Middleton FA, Strick PL (2000) Basal ganglia and cerebellar loops: moto and cognitive circuits. Brain Res Rev 2–3:236–250.

Moon Y (2000) Intimate exchanges: Using computers to elicit self-disclosure from consumers. J Consum Res 26:323–339.

Mori M (1970) The Uncanny Valley (translated by MacDorman KF and Minato T). Energy 7:33–35.

Nass C, Fogg BJ, Moon Y (1996) Can computers be teammates? Int J Hum Comput Stud 45:669–678.

Nass C, Isbister K, Lee E-J (2000) Truth is beauty: Researching embodied conversational agents. Embodied Conversational Agents:374–402.

Nass C, Moon Y (2000) Machines and Mindlessness: Social Responses to Computers. J Soc Issues 56:81–103

Nass C, Moon Y, Carney P (1999) Are respondents polite to computers? Social desirability and direct responses to computers. J Appl Soc Psychol 29:1093–1110.

Natale M (1975) Convergence of mean vocal intensity in dyadic communication as a function of social desirability. J Pers Soc Psychol 32:790–804.

Navailles S, De Deurwaerdère P (2011) Presynaptic control of serotonin on striatal dopamine function. Psychopharmacology (Berl) 213:213–242

Nemeth D, Janacsek K, Csifcsak G, Szvoboda G, Howard Jr. JH, Howard D V. (2011) Interference between sentence processing and probabilistic implicit sequence learning. PloS one2 6.

Norman DA (1981) Categorization of Action Slips. Psychol Rev 88.

Oksama L, Hyönä J (2004) Is multiple object tracking carried out automatically by an early vision mechanism independent of higher-order cognition? An individual difference approach.

Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput Intell Neurosci 2011.

Oslin D, Atkinson RM, Smith DM, Hendrie H (1998) Alcohol related dementia: Proposed clinical criteria. Int J Geriatr Psychiatry 13:203–212.

Oudman E, Van der Stigchel S, Wester AJ, Kessels RPC, Postma A (2011) Intact memory for implicit contextual information in Korsakoff's amnesia. Neuropsychologia 49:2848–2855

Packard MG (2009) Exhumed from thought: Basal ganglia and response learning in the plus-maze. Behav Brain Res:24–31.

Pearson J, Branigan HP, Pickering MJ, Nass CI (2006) Adaptive Language Behavior in HCI: How Expectations and Beliefs about a System Affect Users ' Word Choice. :1177–1180.

Pena J, Hancock JT, Merola N a. (2009) The Priming Effects of Avatars in Virtual Settings. Communic Res 36:838–856.

Petersen SE, Posner MI (2012) The Attention System of the Human Brain: 20 Years After. Annu Rev Neurosci 35:73–89.

Pickering MJ, Branigan HP (1998) The Representation of Verbs: Evidence from Syntactic Priming in Language Production. J Mem Lang 39:633–651.

Pickering MJ, Garrod S (2004) Toward a mechanistic psychology of dialogue. Behav Brain Sci 27:169–225.

Pickering MJ, Garrod S (2013) An integrated theory of language production and comprehension. :329–392.

Pitel A, Berre A Le, Eustache F (2014) Macrostructural abnormalities in Korsakoff syndrome compared with uncomplicated alcoholism.

Poldrack RA, Packard MG (2003) Competition among multiple memory systems : converging evidence from animal and human brain studies. 1497:1–7.

Postle BR, D'Esposito M, Corkin S (2005) Effects of verbal and nonverbal interference on spatial and object visual working memory. Mem Cognit 33:203–212.

Pylyshyn ZW, Storm RW (1988) Tracking multiple independent targets: evidence for a parallel tracking mechanism. Spat Vis 3:179–197

Qu C, Ling Y, Heynderickx I, Brinkman W (2015) Virtual Bystanders in a Language Lesson : Examining the Effect of Social Evaluation , Vicarious Experience , Cognitive Consistency and Praising on Students ' Beliefs , Self- Efficacy and Anxiety in a Virtual Reality Environment. PLoS One 10:1–26.

R Core Development Team (2011) R: A language and environment for statistical computing. Vienna, Austria.

Raz N, Rodrigue KM, Kennedy KM, Head D, Gunning-Dixon F, Acker JD (2003) Differential aging of the human striatum: longitudinal evidence. Am J Neuroradiol:1849–1856.

Rehm M, André E (2005) From Chatterbots to Natural Interaction — Face to Face Communication with Embodied Conversational Agents. IEICE Trans Inf Syst 88:2445–2452.

Reitter D (2008) Context Effects in Language Production : Models of Syntactic Priming in Dialogue Corpora.

Reitter D, Keller F, Moore JD (2011) A Computational Cognitive Model of Syntactic Priming. Cogn Sci 35:587–637.

Rieckmann A, Bäckman L (2009) Implicit learning in aging: extant patterns and new directions. Neuropsychol Rev 19:490–503

Rihs TA, Michel CM, Thut G (2009) NeuroImage A bias for posterior α -band power suppression versus enhancement during shifting versus maintenance of spatial attention. Neuroimage 44:190–199

Rizzo A "Skip", Kim GJ (2005) A SWOT Analysis of the Field of Virtual Reality Rehabilitation and Therapy. Presence Teleoperators Virtual Environ 14:119–146

Rosenthal-von der Pütten AM, Krämer NC, Hoffmann L, Sobieraj S, Eimler SC (2013) An Experimental Study on Emotional Reactions Towards a Robot. Int J Soc Robot 5:17–34.

Salthouse TA (1996) The Processing-Speed Theory of Adult Age Differences in Cognition. Psychol Rev 103:403–428.

Schacter DL, Tulving E (1994) Memory systems. Cambridge, MA: MIT Press.

Scheepers C (2003) Syntactic priming of relative clause attachments:

Persistence of structural configuration in sentence production. Cognition 89:179–205.

Schenkein J (1980) A taxonomy for repeating action sequences in natural conversation. Lang Prod 1:21–47.

Schmand B, Lindeboom J, van Jarskamp F (1992) Manual for the Dutch reading test for adults. Lisse, The Netherlands: Swets & Zeitlinger.

Scholl BJ (2008) 2 What Have We Learned about Attention from Multiple-Object Tracking (and Vice Versa)? Comput Cogn Pylyshyn:49–78

Schoot L, Hagoort P, Segaert K (2014) Bidirectional syntactic priming in conversation: I am primed by you if you are primed by me. In: 20th Annual Architectures and Mechanisms for Language Processing (AMLaP) Conference, pp 57. Edinburgh, Scotland.

Schoot L, Heyselaar E, Hagoort P, Segaert K (2016) Does Syntactic Alignment Effectively Influence How Speakers Are Perceived by Their Conversation Partner? PLoS One 11:e0153521

Schugens MM, Daum I, Spindler M, Birbaumer N (2007) Differential effects of aging on explicit and implicit memory. Aging, Neuropsychol Cogn 4:33–44

Segaert K, Menenti L, Weber K, Hagoort P (2011) A paradox of syntactic priming: why response tendencies show priming for passives, and response latencies show priming for actives. PLoS One 6.

Segaert K, Menenti L, Weber K, Petersson KM, Hagoort P (2012) Shared syntax in language production and language comprehension--an FMRI study. Cereb Cortex 22:1662–1670.

Segaert K, Wheeldon L, Hagoort P (2016) Unifying structural priming effects on syntactic choices and timing of sentence generation. J Mem Lang.

Short E, Hart J, Vu M, Scassellati B (2010) No Fair!! An Interaction with a Cheating Robot. 5th ACM/IEEE Int Conf Human-Robot Interact:219–226

Siegel M, Donner TH, Engel AK (2012) Spectral fingerprints of large-scale neuronal interactions. Nat Rev Neurosci 13:20–25.

Snijders TM, Vosse T, Kempen G, Van Berkum JJ a, Petersson KM, Hagoort P (2009) Retrieval and unification of syntactic structure in sentence comprehension: An fMRI study using word-category ambiguity. Cereb Cortex 19:1493–1503.

Squire LR, Clark RE, Knowlton BJ (2001) Retrograde amnesia. Hippocampus

R

1:50–55.

Squire LR, Dede AJO (2015) Conscious and Unconscious Memory Systems.

Squire LR, Knowlton BJ (2000) The medial temporal lobe, the hippocampus, and the memory systems of the brain. In: The new cognitive neurosciences (Gazzaniga MS, ed), pp 765–780. Cambridge, MA: MIT Press.

Squire LR, Zola SM (1996) Structure and function of declarative and nondeclarative memory systems. Proc Natl Acad Sci:13515–13522.

Stivers T, Enfield NJ, Brown P, Englert C, Hayashi M, Heinemann T, Hoymann G, Rossano F, Peter J, Ruiter D, Yoon K, Levinson SC (2009) Universals and cultural variation in turn-taking in conversation. 106:10587–10592.

Stoyanchev S, Stent A (2009) Lexical and Syntactic Priming and Their Impact in Deployed Spoken Dialog Systems. In: Proceedings of the Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Companion Volume: Short Papers (Ostendorf M, ed), pp 189–192. Boulder, Colorado.

Strayer DL, Watson JM, Drews FA (2011) Cognitive distraction while multitasking in the automobile. Psychol Learn Motiv - Adv Res Theory 54:29.

Street RL, Brady RM, Putman WB (1983) The influence of speech rate stereotypes and rate similarity or listeners' evaluations of speakers. J Lang Soc Psychol 2:37–56.

Sung JE (2015) Age-related changes in sentence production abilities and their relation to working-memory capacity: evidence from a verb-final language. PLoS One 10:e0119424

Sung JE (2016) Age-related Decline in Case-Marker Processing and its Relation to Working Memory Capacity. Journals Gerontol Psychol Sci 0:1–8.

Suzuki WA, Eichenbaum H (2000) The neurophysiology of memory. Ann N Y Acad Sci:175–191.

Swallow KM, Jiang Y V. (2012) Goal-relevant events need not be rare to boost memory for concurrent images. Attention, Perception, Psychophys 74:70–82.

Swallow KM, Jiang Y V. (2013) Attentional load and attentional boost: A review of data and theory. Front Psychol 4:1–13.

Swallow KM, Jiang Y V (2011) The attentional boost effect: Transient increases in attention to one task enhance performance in a second task. Cognition 115:118–132.

Tarr MJ, Warren WH (2002) Virtual reality in behavioral neuroscience and beyond. Nat Neurosci 5 Suppl:1089–1092

Tesink CMJY, Petersson KM, Berkum JJA Van, Buitelaar JK, Hagoort P (2009) Unification of Speaker and Meaning in Language Comprehension: An fMRI Study. J Cogn Neurosci 22:2085–2099.

Tinwell A, Grimshaw M, Williams A (2010) Uncanny behaviour in survival horror games. Games Comput Creat Technol J Artic Paper 13.

Ullman D, Leite I, Phillips J, Kim-Cohen J, Scassellati B (2014) Smart Human , Smarter Robot: How Cheating Affects Perceptions of Social Agency. In: 36th Annual Cognitive Science Society Conference. Quebec City, QB.

Ullman MT (2001) A neurocognitive perspective on language: the declarative/ procedural model. Nat Rev Neurosci 2:717–726.

Ullman MT, Corkin S, Coppola M, Hickok G, Growdon JH, Koroshetz WJ, Pinker S (1997) A Neural Dissociation within Language: Evidence that the Mental Dictionary Is Part of Declarative Memory, and that Grammatical Rules Are Processed by the Procedural System. J Cogn Neurosci 9:266–276.

Ullman MT, Pullman MY (2015) A compensatory role for declarative memory in neurodevelopmental disorders. Neurosci Biobehav Rev 51:205–222

van Diepen RM, Cohen MX, Denys D, Mazaheri A (2015) Attention and Temporal Expectations Modulate Power, Not Phase, of Ongoing Alpha Oscillations. J Cogn Neurosci 27:1573–1586.

Verhage F (1964) Intelligentie en leeftijd: Onderzoek bij Nederlanders van twaalf tot zevenenzeventig jaar. Assen: Van Gorcum.

Vitevitch MS, Sereno J, Jongman A, Goldstein R (2013) Speaker sex influences processing of grammatical gender. PLoS One 8:8–11.

Vivanti G, Hamilton A (2014) Imitation in autism spectrum disorders. … Pervasive Dev Disord …:278–302

Von Humboldt W (1829) On the dual. Nouv Rev Ger 1:378–381.

Vosse T, Kempen G (2000) Syntactic structure assembly in human parsing: a computational model based on competitive inhibition and a lexicalist grammar. Cognition 2:105–143.

Weatherholtz K, Campbell-Kibler K, Jaeger TF (2014) Socially-mediated syntactic alignment. Lang Var Change 26:387–420.

Wester AJ, Van Herten JC, Egger JIM, Kessels RPC (2013) Applicability of the Rivermead Behavioural Memory Test - Third Edition (RBMT-3) in Korsakoff's syndrome and chronic alcoholics. Neuropsychiatr Dis Treat 9:875–881.

Wheeldon L, Smith M (2003) Phrase structure priming: A short-lived effect. Lang Cogn Process 18:431–442

Whittinghill DM, Ziegler B, Moore J, Case T (2015) Nasum Virtualis: A Simple Technique for Reducing Simulator Sickness in Head Mounted VR. In: Game Developers Conference. San Francisco.

Wickens CD (1980) The structure of attentional resources. Atten Perform VIII 8.

Wicklund RA (1975) Objected self-awareness. Adv Exp Soc Psychol 8:233–275.

Willingham DB, Preuss L (1995) The death of implicit memory. Psyche (Stuttg):1–10.

Wilson BA, Greenfield E, Clare L, Baddeley A, Cockburn J, Watson P, Tate R, Sopena S, Nannery R, Crawford J (2008) Rivermead Behavioural Memory Test - Third Edition. Pearson.

Wundt W (1900) Die Sprache. Leipzig: Enghelman.

Yamagishi N, Callen DE, Goda N, Anderson SJ, Yoshida Y, Kawato M (2003) Attentional modulation of oscillatory activity in human visual cortex. Neuroimage 20:98–113.

Zhou FM, Liang Y, Salas R, Zhang L, De Biasi M, Dani JA (2005) Corelease of dopamine and serotonin from striatal dopamine terminals. Neuron 1:65–74.

N

Nederlandse samenvatting

Wanneer we in gesprek zijn met een ander, wordt ons gedrag beïnvloed door onze gesprekspartner. Dit zijn meestal subtiele veranderingen waar we ons niet eens bewust van zijn. Bijvoorbeeld, als we een gesprek voeren met iemand die erg snel, of juist erg langzaam praat, gaan we zelf ook ietsje sneller (of langzamer) praten dan we normaal zouden doen. Dit soort aanpassingen aan onze gesprekspartner zien we niet alleen terug in de snelheid waarmee we spreken, maar ook in onze woordkeus, het gebruik van gebaren tijdens een gesprek, enzovoorts. In dit proefschrift kijk ik specifiek naar de opbouw van de zinstructuur. Als spreker één een plaatje beschrijft met een passieve zin, zoals de man wordt door de vrouw gekust, dan is spreker twee meer geneigd om het volgende plaatje ook te beschrijven met een passieve zin, zoals; het meisje wordt door de jongen geknuffeld (in plaats van een actieve zin zoals de jongen knuffelt het meisje). In de wetenschap (en in dit proefschrift) wordt dit het syntactic priming effect genoemd: we nemen de voorkeuren in zinsstructuur (syntax) over van onze gesprekspartner. Het syntactic priming effect is op zichzelf al interessant, omdat het ons vertelt dat begrip en productie van zinstructuur sterk aan elkaar gerelateerd zijn. Het is verder interessant omdat zinstructuur vrij abstract is: de twee voorbeeldzinnen hierboven, bijvoorbeeld, (de man wordt door de vrouw gekust en het meisje wordt door de jongen geknuffeld) hebben weinig met elkaar gemeen, alleen de abstracte woordvolgorde (maar niet de inhoudswoorden zelf) wordt herhaald.

Omdat de spreker zich meestal niet bewust is dat zijn/haar keuze voor een bepaalde zinsstructuur wordt beïnvloed door zijn/haar gesprekspartner, zijn er veel theorieën die stellen dat syntactic priming onz impliciete geheugen gebruikt. Dat betekent dat we syntactische informatie opslaan in het deel van ons geheugen zonder dat we daar bewust van zijn. Maar juist omdat ons impliciete geheugen onbewust is, is het moeilijk om hiermee te experimenteren zonder dat proefpersonen zich er bewust van worden, en daardoor hun expliciete, bewuste geheugen gaat gebruiken. In hoofdstuk 2 kijk ik daarom of patiënten die geen gebruik meer kunnen maken van hun bewuste geheugen ook een syntactic priming effect laten zien. Als dat zo is, is dit sterke evidentie dat syntactic priming effecten gedreven worden door het onbewuste geheugen, omdat hun bewuste geheugen niet meer werkt. In deze studie meet ik 17 patiënten,. De patiënten in mijn studie werden sterk beïnvloed door de syntactische keuzes van hun gesprekspartner (syntax priming effect), waarmee wordt bewezen dat syntactic priming inderdaad gedreven wordt door ons onbewuste geheugen. Een interessant maar onverwacht resultaat is dat de gezonde proefpersonen in deze studie geen syntactic priming effect lieten zien. Met een gemiddelde leeftijd van 62 jaar waren de proefpersonen (en patiënten) in dit onderzoek ouder dan de studenten die we meestal meten in dit soort studies. Eén van de vragen die dus

N

overblijft na afloop van hooftstuk 2 is of syntactic priming, het overnemen van de voorkeur van onze gesprekspartner voor een bepaalde zinstructuur, iets is dat we niet meer kunnen doen na een bepaalde leeftijd, en wat voor impact dat zou kunnen hebben op hoe we met anderen communiceren.

Verder heb ik gefocust op hoe syntactic priming effecten beïnvloed worden door wat we denken van onze gesprekspartner. Er zijn veel theorieën die stellen dat syntactic priming een sociaal doel heeft; we zouden de voorkeuren in zinsstructuur van onze gesprekspartner overnemen omdat we affiniteit willen overbrengen. Theoretisch klinkt deze hypothese logisch, maar er zijn maar enkele studies die deze hypothese daadwerkelijk hebben onderzocht. Bovendien tonen deze studies tegenoverstelde resultaten: de ene studie zegt dat sprekers een groter syntactic priming effect laten zien (dus hun gesprekspartner meer nadoen) als ze hem of haar leuk vinden, de ander zegt dat sprekers juist een kleiner syntactic priming effect laten zien wanneer ze hun gesprekspartner leuk vinden. In hoofdstukken 3 en 4 doe ik mijn eigen onderzoek naar deze vraag. Ik gebruik hiervoor Virtual Reality, wat betekent dat de proefpersonen een gesprek voeren met een digitaal persoon (een “avatar”). Door een avatar te gebruiken kan ik er zeker van zijn dat de gesprekspartner met wie de proefpersonen praten tijdens het experiment zich altijd hetzelfde gedraagt. Omdat ik meer dan 50 proefpersonen test in hoofdstuk 4, zouden de resultaten heel anders zijn als ik mijn proefpersonen zou hebben laten praten met een echt mens: bij proefpersoon 50 zou de gesprekspartner er misschien wel genoeg van hebben om dezelfde taak voor de 50ste keer te doen. Maar eerst moet ik laten zien dat het syntactic priming effect van proefpersonen hetzelfde zijn bij zowel de avatar als een echt persoon. In hoofdstuk 3 laat ik zien dat het gedrag hetzelfde is, en dus kunnen we (voor deze taak) het gesprekspartner inwisselen voor een avatar. In hoofdstuk 4 laat ik zien dat wat de proefpersonen van de avatars denken hun syntactic priming effect beinvloedt, maar het effect is minder rechtlijnig dan verwacht. Proefpersonen hebben het sterkste syntactic priming effect (dus herhalen de zinsstructuren van de avatar het meeste) als ze de avatar best oké vinden. Als ze de avatar helemaal geweldig of juist echt niet geweldig vinden, zien we een kleiner syntactic priming effect.

In de laatste hoofdstukken van dit proefschrift kijk ik waar dit effect vandaan komt. Waarom doen we een persoon die we best oké vinden meer na dan iemand die we geweldig of juist niet geweldig vinden? Mijn theorie is dat het met aandacht te maken heeft. Als we iemand geweldig of juist niet geweldig vinden, zijn we misschien afgeleid door de eigenschappen van die persoon en letten daarom minder op hoe deze spreekt. Daarom kijk ik in hoofdstukken 5 en 6 of de sterkte van de syntactic priming effect inderdaad met aandacht te maken heeft. In hoofdstuk 5 doen de proefpersonen een syntactic priming taak terwijl ze tegelijkertijd een tweede taak doen. Deze tweede taak leidt niet, een beetje, of juist heel erg af. De resultaten laten inderdaad zien dat als de proefpersonen niet, of heel erg afgeleid zijn, ze het kleinste syntactic priming effect laten zien. Waarom ze een kleiner syntactic priming effect laten zien als ze juist minder afgeleid zijn is een vraag die ik op dit moment helaas niet kan beantwoorden.

In hoofdstuk 6 link ik het concept afleiding aan de opinie van de avatar via de beeldvormingstechniek electroencephalographie (EEG). Via EEG kan ik de hersenactiviteit van proefpersonen meten terwijl ze naar een avatar kijken. Ik laat zien dat de proefpersonen het meeste aandacht geven aan de avatar die ze best oké vinden, en het minste aan de avatar die ze geweldig of juist niet geweldig vinden. Als je deze resultaten combineert met de resultaten van hoofdstuk 4 en 5 is het duidelijk dat als je meer aandacht hebt voor je gesprekspartner, je ook hun taalgebruik sneller overneemt. Als je minder aandacht hebt voor je gesprekspartner omdat je ze geweldig of juist niet geweldig vindt, dan let je ook minder op hun taalgebruik en verkleint dit je syntactic priming effect.

Samenvattend laten de resultaten in mijn proefschrift zien dat veranderingen in hoeveel we de voorkeuren van onze gesprekspartner voor een bepaalde zinstructuur overnemen niet zo zeer te maken hebben met hoe goed we de zinstructuur begrijpen en verwerken in ons brein. In plaats daarvan concludeer ik dat het meer te maken heeft met ons onbewuste geheugen en aandacht: hoe meer aandacht we aan onz gesprekspartner geven, hoe beter we onbewust hun voorkeur voor een bepaalde zinstructuur opslaan. Dit beïnvloedt de waarschijnlijkheid dat we die zinstructuur zelf gaan gebruiken.

C

Curriculum Vitae

Evelien left Haarlem when she was 5 years old. Her family first moved to Accra, Ghana, where she learned English and discovered her love for fried plantain. At the age of 11, she moved (with her heavy Ghanian accent) to Texas, where she learnt to pledge her allegiance to the USA flag twice a day everyday, and that hiding under a desk would save you from the hurricane that never arrived (although, years later, she did finally meet her first hurricane by sailing through it – it wasn't as intense as her teacher made it out to be). After a year of the Lone Star State, she moved to Surrey, England where she switched her now Texan accent for a British one, and was introduced to Cadbury's chocolate and Jaffa Cakes. At the age of 15 she moved to Abu Dhabi (to a British School so no change of accent here) where she discovered her love of Lebanese food, especially Middle Eastern Shwarma (with chicken, pickles and French fries inside the wrap – the way it is supposed to be!). After four years it was finally time to go to university, so naturally she moved to Canada, where she switched her 6 year old British accent for a Canadian one – the one that she still has to this day. In Canada she discovered her love for Beaver Tails and, most importantly, discovered Neuroscience. She completed a Life Sciences Bachelors with a specialization in Neuroscience (there were no neuroscience majors at her university yet) at Queen's University in Kingston (the first capital of Canada!) completing a thesis on training macaque monkeys a visual working memory task. She stayed in the lab, with her trained monkeys, to complete a Masters of Science investigating the neural basis of working memory capacity. She then took a year off to sail around the world, eventually ending up in the Netherlands by pure coincidence, where she managed to wiggle her way into a PhD position with Peter Hagoort and Katrien Segaert at the Max Planck Institute. I use the word wiggle here, because she was originally not hired for the PhD position she applied for and instead worked as an intern for 6 months while she developed the VR-based PhD project you are holding in your hands. It is now 3 years later and she has a PhD, a husband, a car, five chickens, two cats, and a dog.

Publications

Nieuwland M, Politzer-Ahles S, Heyselaar E, Segaert K, Darley E, Kazanina N, ... & Mézière, D (2017) Limits on prediction in language comprehension: A multi-lab failure to replicate evidence for probabilistic pre-activation of phonology. *bioRxiv*, 111807.

Heyselaar E, Hagoort P, Segaert K (2017) How social opinion influences syntactic processing - an investigation using Virtual Reality. *PLoS One* 12(4), e0174405.

Schoot L, Heyselaar E, Segaert K, Hagoort P (2016) Syntactic alignment in conversation: an effective way of getting your partner to like you? *PLoS One* 11(4), e0153521.

Heyselaar E, Hagoort P, Segaert K (2015) In dialogue with an avatar, language bheaviour is identical compared to dialogue with a human partner. *Behavior Research Methods* 49(1), 46-60.

Heyselaar E, Hagoort P, Segaert K (2014) In dialogue with an avatar, syntax production is identical compared to dialogue with a human partner. In The 36th annual meeting of the Cognitive Science Society (CogSci 2014) (pp. 2351 - 2356). *Cognitive Science Society*.

Heyselaar E (2011) Role of NMDA in the visual working memory of the macaque monkey. Master's Thesis; *Queen's University Archives*

Heyselaar E, Johnson K, Pare M (2011) A change detection approach to study visual working memory of the macaque monkey. *Journal of Vision* 11(3).

A

Acknowledgements

Most importantly, I want to thank my parents. I had applied to many PhD positions in Europe and was happily invited to interview for two at the Neurobiology of Language department with Roel Willems and/or Dan Acheson as daily supervisors. This was the only interview offer I received and, you'll never guess, I was rejected for both. I was at a loss. I had no other prospects and by this time it was October and hiring season was more or less over. My parents convinced me to offer my "research services" for free, to get my foot in the door and make them realize they had made a mistake. Even though I had been financially independent for many years, they offered to go back to financially supporting me. So from November 2012 until July 2013 I worked part time getting my virtual reality project off the ground, the rest of the day working at a call center for electrical lease cars (Car2Go). Without my parents, there very literally would have been no thesis.

Of course, they also get a very big thank you for always wanting to read all my papers once they had been written, and would smile and nod enthusiastically when we talked about it, and then some months later admit that they had no idea what they were reading, but it sounded fancy so it must be very important.

Next up is Katrien Segaert. Because without her there also would have been a lot less of a thesis. Katrien, thank you for allowing me to bother you relentlessly during my internship days, so much so that one day we looked up and realized that you were basically my daily supervisor. And then you were stuck with me for 3 additional years ;) Thank you for encouraging freedom and creativity – I don't know if you realize how rare of a person you are in letting Lotte and I more or less rein free. The trust and encouragement you gave has changed me so much, and I am definitely a much better scientist and a much kinder person because of you. May we have many more years of me chasing cats off of your cupboards.

And then there is Peter Hagoort. Thank you for allowing me to camp out in an empty office for free and work to prove that I deserved to complete a PhD with you. I still remember that day in June when I asked you if I could start a PhD for a second time. I was so nervous, so afraid that you would realize that I basically had a vague idea based on an article my dad had sent me and that I could also just have completed it as a research assistant (thanks Katrien for putting that idea in my head 5 minutes before my meeting). I had practised every outcome in my head and when it was finally time, I walking into your office, you looked up and simply said, "so you want to do a PhD? When do you want to start?" before I could even say anything about why I was in your office.

Thank you to my manuscript committee for reading all of it, finding typos, giving new analysis suggestions, and also giving suggestions on how to publish

the chapters that are not yet published. I was surprised by the depth of feedback I received, and I feel my thesis and my future publications have benefited greatly from you taking the time to write up detailed changes. Thank you also to my corona, whoever you end up being, for taking time out of your busy schedule to read the thesis and come up with interesting questions to ask. Thank you also for opting to spend some hours in Nijmegen listening to me prattle on about why I believe in the work I did for the past three years.

Before this acknowledgments section disintegrates into me naming people at random hoping that I don't forget anyone, I need to thank Lotte and Gwilym. Lotte, on my very first day at the MPI (first day of my internship), you came by and asked if I wanted to go to lunch. Since then, there are so many fun memories that I can't really express how grateful I am to have you as a friend. From confusing eavesdropping train goers with our Dutch/English verbal mix, to the Christmas Horse, stoel!, and visiting that town Bham together (which I think is near Birmingham but I'm not sure). Gwilym, at first I was a bit wary of you, but after we connected on our love of memes, all things stupid internet, and long intense conversations about bullshit, I'm surprised we got any work done at all! You two shaped my experience at the MPI.

Thank you to Jobke, for brainstorming solutions when I was stuck, for letting me ramble on about projects when you weren't even sure what my thesis was actually about, for letting me convince you to make films for the directors, and for taking care of me when I was just sick of it all. If you weren't sure why I wanted you to be my paranymph, hopefully it's a bit more obvious now.

And of course a super special thank you to Amie, for being you and for rearranging your schedule so that you can be my paranymph ☺ I'm not entirely sure how we met, but you always make me feel super happy and upbeat ☺

And now comes the time when I randomly say thanks to a bunch of people and hope that I don't forget anyone really important. If you are not in here, that doesn't mean I don't care! I means that I shouldn't have written this the night before my thesis had to go to print. Please know that I still care.

Thank you so much to Nadine de Rue, Charlotte Poulisse, Laura Arendsen, and Vera van 't Hoff for helping me collect copious amounts of data, or being the confederate over 50 times!

Thank you to the NBL department for creating such a relaxed yet encouraging work environment. This department was very different to the one I knew in Canada, and at first I was a bit intimidated. I think few PhD students know how lucky they are to work in such a supportive and gezellige lab. I wanted to do a "special shout-out to" but I just ended up listing pretty much everyone in our lab ☺

Thank you to everyone else at the MPI, for happy greetings, interesting conversations, and endless amounts of support (especially you, Dirkje van der Aa, Kevin Lam, and Karin Kastens!).

Thank you also to Bianca Weiland and Ilona Berendsen-Weerkamp, for the fun times we had together when I was working at the call center, and for being the first pilot participants to basically all my experiments!

Thank you to Linda Wheeldon, Ali Mazaheri, Roy Kessels, Florian Jaeger, and Kodi Weatherholtz for brainstorming, giving advice on my projects, or agreeing to co-author a project with me!

And thank you to the School of Psychology at the University of Birmingham for letting me present most of my projects at the lab meetings for feedback and ideas, and letting me terrorize the local PhD students multiple times per year for the past two years!

This thesis was built on 780 cups of coffee, 427 participants, 0 burn outs, and lots of fun times.

M

MPI Series in Psycholinguistics

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

21. Fixed expressions and the production of idioms. *Simone A. Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure: Distinguishing contrastive from

non-contrastive variation. *Christiane Dietrich*

41. Cognitive cladistics and the relativity of spatial cognition. *Daniel B.M. Haun*

42. The acquisition of auditory categories. *Martijn Goudbeek*

43. Affix reduction in spoken Dutch. *Mark Pluymaekers*

44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*

45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*

46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*

47. Crosslinguistic influence in first and second languages: Convergence in speech and gesture. *Amanda Brown*

48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*

49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*

50. Lexical processing of morphologically complex words: An information-theoretical perspective. *Victor Kuperman*

51. A grammar of Savosavo, a Papuan language of the Solomon Islands. *Claudia Wegener*

52. Prosodic structure in speech production and perception. *Claudia Kuzla*

53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*

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

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

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

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

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

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

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

61. Split intransitivity in Rotokas, a Papuan language of Bougainville. *Stuart Robinson*
62. Evidentiality and intersubjectivity in Yurakaré: An interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension: A neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*
65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*
66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*
67. The relevance of early word recognition: Insights from the infant brain. *Caroline Junge*
68. Adjusting to different speakers: Extrinsic normalization in vowel perception. *Matthias J. Sjerps*
69. Structuring language. Contributions to the neurocognition of syntax. *Katrien R. Segaert*
70. Infants' appreciation of others' mental states in prelinguistic communication: A second person approach to mindreading. *Birgit Knudsen*
71. Gaze behavior in face-to-face interaction. *Federico Rossano*
72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space. *Conny de Vos*
73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*
74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt Witteman*
75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*
76. Territories of knowledge in Japanese conversation. *Kaoru Hayano*
77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder*
78. Contributions of executive control to individual differences in word production. *Zeshu Shao*
79. Hearing speech and seeing speech: Perceptual adjustments in auditory-visual processing. *Patrick van der Zande*

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

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

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

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

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

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

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

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

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

89. Rediscovering a forgotten language. *Jiyoun Choi*

90. The road to native listening: Language-general perception, language-specific input. *Sho Tsuji*

91. Infants'understanding of communication as participants and observers. *Gudmundur Bjarki Thorgrímsson*

92. Information structure in Avatime. *Saskia van Putten*

93. Switch reference in Whitesands. *Jeremy Hammond*

94. Machine learning for gesture recognition from videos. *Binyam Gebrekidan Gebre*

95. Acquisition of spatial language by signing and speaking children: a comparison of Turkish sign language (TID) and Turkish. *Beyza Sümer*

96. An ear for pitch: on the effects of experience and aptitude in processing pitch in language and music. *Salomi Savvatia Asaridou*

97. lncrementality and Flexibility in Sentence Production. *Maartje van de Velde*

98. Social learning dynamics in chimpanzees: Reflections on (nonhuman) animal culture. *Edwin van Leeuwen*

99. The request system in Italian interaction. *Giovanni Rossi*

100. Timing turns in conversation: A temporal preparation account. *Lilla Magyari*

101. Assessing birth language memory in young adoptees. *Wencui Zhou*

102. A social and neurobiological approach to pointing in speech and gesture. *David Peeters*

103. Investigating the genetic basis of reading and language skills. *Alessandro Gialluisi*

104. Conversation Electrified: The Electrophysiology of Spoken Speech Act Recognition. *Rósa Signý Gisladottir*

105. Modelling Multimodal Language Processing. *Alastair Smith*

106. Predicting language in different contexts: The nature and limits of mechanisms in anticipatory language processing. *Florian Hintz*

107. Situational variation in non-native communication. *Huib Kouwenhoven*

108. Sustained attention in language production. *Suzanne Jongman*

109. Acoustic reduction in spoken-word processing: Distributional, syntactic, morphosyntatic, and orthographic effects. *Malte Viebahn*

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

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

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

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

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

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

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

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

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

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

Manrique

120. Talking Sense: the behavioural and neural correlates of sound symbolism. *Gwilym Lockwood*

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

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

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

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