Explorations of Beta-band Neural Oscillations During Language Comprehension

Ashley Lewis

MPI Series

ISBN: 978-90-76203-85-0

123

In the end the party would announce that two and two made five, and you would have to believe it.

It was inevitable that they should make that claim sooner or later; the logic of their position demanded it.

Not merely the validity of experience, but the very existence of external reality was tacitly denied by their philosophy.

Explorations of Beta-band Neural Oscillations During Language Comprehension: Sentence Processing and Beyond

Ashley Glen Lewis

Max Planck Institute for Psycholinguistics

Series

Invitation

You are hereby invited to the defense of my PhD thesis entitled

Explorations of Beta-band Neural Oscillations During Language Comprehension

On Monday, May 8, 2017 at 14:30 in the aula of the Radboud University Nijmegen, Comeniuslaan 2.

After the defense there will be a reception at the Max-Planck Institute, Wundtlaan 1.

Ashley Lewis
lewis@haskins.yale.edu

Paranymphs

Suzanne Jongman
suzanne.jongman@mpi.nl

Bart Gips
bart.gips@gmail.com

Max Planck Institute for Psycholinguistics

Series

Explorations of Beta-band Neural Oscillations During Language Comprehension: Sentence Processing and Beyond

Cover Text: Quoted from Nineteen Eighty-Four by George Orwell

ISBN: 978-90-76203-85-0

Printed and bound by Ipskamp Drukkers b.v.

The research reported in this thesis was supported by an IMPRS fellowship funded by the Max Planck Gesellschaft, München, Germany.

Explorations of beta-band neural oscillations during language comprehension:
Sentence processing and beyond

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 maandag 8 mei 2017
om 14.30 uur precies

door
Ashley Glen Lewis
geboren op 8 Januari 1985
te Durban (Zuid Afrika)

Promotor:

Prof. dr. H.J. Schriefers

Copromotoren:

Dr. M.C.M. Bastiaansen (NHTV Breda)

Dr. J. Schoffelen

Manuscriptcommissie:

Prof. dr. P. Hagoort

Prof. dr. O. Jensen (University of Birmingham, Verenigd Koninkrijk)

Prof. dr. D. Roehm (Universität Salzburg, Oostenrijk)

Explorations of beta-band neural oscillations during language comprehension:
Sentence processing and beyond

Doctoral thesis

to obtain the degree of doctor

from Radboud University Nijmegen

on the authority of the rector magnificus, prof. dr. J.H.M. van Krieken

according to the decision of the council of deans

to be defined in public on monday may 8, 2017

at 14.30 hours

by

Ashley Glen Lewis

born on january 8, 1985

in Durban (South Africa)

Supervisor:

Prof. dr. H.J. Schriefers

Co-supervisors:

Dr. M.C.M. Bastiaansen (NHTV Breda University of Applied Sciences)

Dr. J. Schoffelen

Manuscript committee:

Prof. dr. P. Hagoort

Prof. dr. O. Jensen (University of Birmingham, United Kingdom)

Prof. dr. D. Roehm (Universität Salzburg, Austria)

I’ll take the haunches. All the rest can be a bonus.

Chapter 1

General Introduction

Partially based on:

Lewis, A. G., Wang, L., & Bastiaansen, M. (2015). Fast oscillatory dynamics during language comprehension: Unification versus maintenance and prediction? *Brain and language*, *148*, 51-63.

Lewis, A. G., & Bastiaansen, M. (2015). A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex*, *68*, 155-168.

Lewis, A. G., Schoffelen, J. M., Schriefers, H., & Bastiaansen, M. (2016). A predictive coding perspective on beta oscillations during sentence-level language comprehension. *Frontiers in human neuroscience*, *10*.

1.1 Some scene-setting

Language comprehension requires the fast and efficient integration of information represented at a multitude of different levels and timescales (Jackendoff, 2007). This means that numerous different and often spatially distant brain regions have to interact quickly and dynamically in order to achieve even the most basic linguistic processing. It is therefore not surprising that oscillatory neural dynamics have been steadily receiving more attention as a robust and temporally precise signature of network activity related to language processing (Friederici & Singer, 2015; Weiss & Mueller, 2012). Recently there has been a great deal of interest in oscillatory activity in the beta frequency range (13 to 30 Hz) as a potential index of syntactic integration/unification during language comprehension (e.g., Bastiaansen & Hagoort, 2015; Bastiaansen, Magyari, & Hagoort, 2010). On the other hand, there is reason to think that a strict link between beta oscillations and syntactic processing does not hold (e.g., Wang, Jensen et al., 2012; see Lewis, Wang, & Bastiaansen, 2015 for discussion), and there are plausible alternative proposals for how beta is related to language processing (Lewis et al., 2015; Weiss & Mueller, 2012). This dissertation explores oscillatory neural dynamics in the beta frequency range related to language comprehension beyond the level of individual words. It furthermore explicitly investigates the extent to which oscillatory neural dynamics in the beta frequency range during language comprehension are specifically an index of syntactic processing, or alternatively reflect network dynamics related to more domain-general processing, which can sometimes be recruited in the service of syntactic processing.

1.1.1 Networks of dynamically interacting brain regions

Neuroimaging methods with high spatial resolution, like functional magnetic resonance imaging (fMRI) and positron emission tomography (PET), have resulted in a detailed specification of the brain regions that are differentially activated in support of various cognitive functions. Perisylvian cortical regions (including left inferior frontal and left temporal cortex) have been identified as critical for supporting the computational machinery associated with various aspects of linguistic processing (e.g., Hagoort, 2013, 2014; Hickok & Poeppel, 2007; see Hagoort & Indefrey, 2014 for a meta-analysis). Furthermore, the emergence of methods like diffusion tensor imaging (DTI), and related methods for non-invasively mapping the white matter tracts of the brain, has improved our understanding of the anatomical connectivity between various regions considered important for linguistic processing (e.g., Catani et al., 2013; Friederici, 2009; Roelofs, 2014). It has become clear that a one-to-one mapping between brain regions and cognitive functions is too simplistic a view to explain how the brain supports cognitive processing (e.g., Fedorenko & Thompson-Schill, 2014). As a result, spatially distributed, yet functionally coherent networks are increasingly being thought of as the most relevant unit of analysis in cognitive neuroscience (e.g., Fox et al., 2005; Sporns, 2012; Sporns, Chialvo, Kaiser, & Hilgetag, 2004; Varela et al., 2001). This is no less the case for language processing, and raises the question of which mechanism(s) are responsible for the dynamic recruitment of the various participating cortical and sub-cortical areas.

1.1.2 Neural oscillations as a window onto network dynamics

A large amount of evidence has accumulated over the last two or more decades suggesting that the coupling and uncoupling of functional networks in the brain is related to patterns of neural synchronization and desynchronization (Bastiaansen, Mazaheri, & Jensen, 2012; Bastiaansen & Hagoort, 2006; Pfurtscheller & Lopes da Silva, 1999; Singer, 1993, 2011; Varela, Lachaux, Rodriguez, & Martinerie, 2001; Womelsdorf et al., 2007). One instance of this occurs when areas that are part of the same functional network are linked by synchronous rhythmic firing in the same frequency range. Conceptually, synchronous repetitive firing of neurons increases the probability that they entrain one another and thereby activates participating functional networks at particular frequencies (König & Schillen, 1991). In this way the brain achieves frequency-specific segregation of information being processed by different functional networks. On the other hand, frequency-specific oscillatory neural synchrony also binds together information represented in different elements or subcomponents of the same functional network (Gray, Konig, Engel, & Singer, 1989). Such oscillatory neural phenomena typically have similar functions across multiple spatial and temporal scales. Modulations of frequency-specific power are often associated with synchrony within local neural populations, while modulations of frequency-specific phase coupling measures (e.g. coherence or phase-locking value) are most often associated with synchrony between more distant neural populations (inter-areal synchrony). There is however no clear distinction between local and inter-areal synchrony, and hence no guarantee that power always measures local synchrony and coherence always measures inter-areal communication

(Varela et al., 2001). The primary focus in this dissertation will be on power measures as an index of frequency-specific changes in oscillatory synchrony.

1.1.3 Measuring neural oscillations

In order to empirically study the role of oscillatory dynamics in functional neural network formation, one needs to address the question of how to quantify the rapidly changing patterns of synchronization and desynchronization of neural activity. First, the fast temporal dynamics involved can only be captured by imaging methods with a high temporal resolution such as EEG and MEG. However, the standard analysis techniques used in EEG and MEG research, which involve the computation of ERPs/ERFs, capture only a restricted range of the underlying neural activity (i.e. evoked activity that is strongly time- and phase-locked to an event of interest, but not induced activity, which exhibits time-, but not phase-locking; see e.g., Bastiaansen et al., 2012, or Makeig, Debener, Onton, & Delorme, 2004 for a more detailed discussion). Rather (restricting ourselves to scalp-recorded activity), two related measures are typically used in this context, namely power and coherence changes (Varela et al., 2001). Event-related changes in frequency band-specific power reflect changes in synchronization of local underlying neural tissue (i.e., within the nodes of a distributed network). Such power changes are typically quantified by means of wavelet analysis (e.g., Tallon-Baudry, Bertrand, Peronnet, & Pernier, 1998) or multitaper analysis (Mitra & Pesaran, 1999). In turn, event-related changes in frequency band-specific (phase) coherence reflect changes in synchronization between (often, but not necessarily, spatially distant) brain areas (i.e., between the nodes of a distributed network; see Bastiaansen & Hagoort, 2006; Bastiaansen et al., 2012).

Power and coherence measures can be seen as complementary to ERPs/ERFs, providing additional information about the underlying neural dynamics that may be overlooked when only using these more traditional approaches. For instance, increases in power are thought to reflect increased synchrony, and hence coupling of the nodes of a functional network. Decreases in power may (at least some of the time) be related to larger processing costs (e.g., alpha power desynchronization is thought to reflect increased attentional demands, cf., Jensen & Mazaheri, 2010). In contrast, larger ERP components are usually related to more effortful processing (for example when unification, or more general processing difficulties arise).

It should be mentioned that besides power and coherence changes, other event-related changes in oscillatory EEG/MEG activity could be meaningful in studying the neural basis of cognitive functions (see Makeig, Debener, Onton, & Delorme, 2004 for an excellent overview of potentially relevant phenomena). For instance, there have been several reports of event-related phase resetting (the phase realignment, over trials, of oscillatory activity with respect to an experimental event; e.g., Braeutigam, Bailey, & Swithenby, 2001; Rizzuto et al., 2003). In addition, oscillatory entrainment (mainly in the delta, theta and gamma frequency ranges) has been linked to the 'packaging' of information on varying timescales during speech perception (e.g. Giraud & Poeppel, 2012). Finally, recently there has been much interest in cross-frequency coupling, where the phase or amplitude of low frequency oscillatory activity modulates the phase or amplitude of oscillatory activity at higher frequencies (e.g., Lisman & Jensen, 2013). In this dissertation I will restrict myself to measures of oscillatory power as an index of frequency-specific changes in neural network dynamics as outlined above.

1.2 The big question

A growing body of literature has accumulated relating sentence-level language comprehension to event-related changes in EEG and MEG oscillations (Bastiaansen & Hagoort, 2015; Bastiaansen, Magyari, & Hagoort, 2010; Peña & Melloni, 2012; for reviews see Bastiaansen et al., 2012; Weiss & Mueller, 2012). Such studies typically investigate patterns of temporal dynamics, that are arguably associated with the coupling and uncoupling of nodes in the brain's language network. Effects have been found in all the classical frequency ranges, with for example theta oscillations (3 to 7 Hz) linked to lexical retrieval operations and semantic working memory, alpha (8 to 12 Hz) linked to task-specific working memory load (Bastiaansen et al., 2012; Bastiaansen & Hagoort, 2006; Weiss et al., 2005), and gamma (30 to 100 Hz and beyond) linked to semantic unification/integration operations during sentence processing (e.g., Bastiaansen & Hagoort, 2015). In this dissertation I will focus on beta oscillations and the role they play in supporting language comprehension.

1.2.1 Beta and syntactic processing

A number of studies have compared syntactically acceptable sentences (e.g., 'Janneke got the blessing at the river') to sentences containing a syntactic violation (e.g., 'Janneke got the to bless at the river'; Bastiaansen et al., 2010; Davidson & Indefrey, 2007; Kielar, Meltzer, Moreno, Alain, & Bialystok, 2014; Kielar, Panamsky, Links, & Meltzer, 2015; Pérez, Molinaro, Mancini, Barraza, & Carreiras, 2012). They have all reported that power in the beta frequency range is higher at the target word for syntactically acceptable sentences compared to sentences containing a syntactic violation. Bastiaansen et al. (2010) have also shown that power in the beta band is higher at the target word for syntactically acceptable sentences compared to the same words in random order (e.g., 'The the Janneke blessing got river at'). Extending these findings, Bastiaansen et al. (2010) showed that beta power increases linearly over the course of syntactically acceptable sentences, and remains consistently low over the course of random word lists (see also Bastiaansen & Hagoort, 2015 for a replication of these findings). For sentences containing syntactic violations beta-power shows a linear increase up to the point of the violating word and then rapidly returns to baseline levels. These findings have been taken as support for the idea that oscillatory activity in the beta frequency range might be related to syntactic unification operations during sentence-level language comprehension.

Further support for this idea comes from studies showing that beta power is higher for sentences which are more demanding in terms of syntactic unification load than for less demanding sentences. Bastiaansen and Hagoort (2006) reported that beta power was higher for syntactically more demanding center-embedded (e.g., 'The juice that the child spilled stained the rug') compared to right-branching relative clauses (e.g., 'The child spilled the juice that stained the rug'). Similarly, syntactically more demanding object-relative clauses showed higher beta coherence just after the relative clause than their simpler subject-relative counterparts (Weiss et al., 2005). Meyer et al. (2013) showed that beta power was higher for long- compared to short-distance subject-verb agreement dependencies at the point in a sentence where the agreement relation between a subject and subsequent verb had to be computed. Since syntactic working memory load is higher for long- compared to short-distance dependencies, this leads to higher load on the system responsible for syntactic unification, and this result can thus be interpreted as support for a link between beta and syntactic unification. For the arguments that will follow it is important to note that none of the above constructions involved locally ambiguous sentences (in Weiss et al, 2005 it was always clear whether the sentence was a subject- or object-relative clause), and so it was never the case that participants had a clear a priori preference for a particular sentence construction type (e.g., subject-relative clause), which was then overridden by the input disambiguating towards an alternative construction type (e.g., object-relative clause).

These findings (together with a number of gamma findings not reviewed here) resulted in the proposal of the 'frequency-based segregation of syntactic and semantic unification hypothesis' (Bastiaansen & Hagoort, 2015). Under this proposal, oscillatory activity in the beta and gamma frequency ranges constitute separate frequency 'channels' along which syntactic and semantic unification operations respectively can operate, without interfering with one another. Importantly, this proposal entails a strong link between oscillatory activity in the beta frequency range and syntactic processing, and that is one of the hypotheses that will be tested in this dissertation. This beta-syntax link suggests that whenever syntactic processing is more demanding (e.g., long- compared to short-distance subject-verb agreement dependencies) beta power should be higher. Conversely, whenever syntactic processing is disrupted (e.g., when encountering a syntactic violation) beta power should be lower.

So far, all the evidence seems to point strongly to a link between oscillatory activity in the beta band and syntactic unification operations during sentence-level language comprehension. Not all the data are consistent with this interpretation however. For one thing, semantic anomalies (e.g., 'The climbers finally reached the top of the *tulip*') elicit decreases in beta power relative to semantically (and syntactically) acceptable sentences (e.g., 'The climbers finally reached the top of the *mountain*'; Kielar et al., 2014, 2015; Luo, Zhang, Feng, & Zhou, 2010; Wang, Jensen et al., 2012). Luo et al. (2010) also showed a beta power decrease for rhythmically abnormal target nouns (in verb-noun pairs in Chinese) compared to their rhythmically normal counterparts. Furthermore, Pérez et al. (2012) showed that for Spanish 'Unagreement' (where there is a mismatch between the grammatical person feature marking on the subject and the verb of a sentence, but where that sentence still remains perfectly grammatical; see Pérez et al., 2012 for more details) there is a decrease in beta power (similar to the beta power decrease reported above for the genuine agreement violation condition in that study) compared to syntactically acceptable sentences. Since 'Unagreement' does not strictly speaking represent a case of syntactic violation it is not clear why syntactic unification should be disrupted in this case, and so the beta power decrease observed there is unlikely to reflect syntactic unification difficulties.

1.2.2 Beta and maintenance/change of the current cognitive set

An alternative hypothesis has been proposed (Lewis & Bastiaansen, 2015; Lewis et al., 2015), suggesting that the more domain-general framework of Engel and Fries (2010) might provide a better explanation for the beta findings reported above. On that account beta power increases reflect active maintenance of the current cognitive set (which in the case of language comprehension has been defined as the current sentence-level meaning representation under construction; Lewis et al., 2015), while decreases in beta power reflect a change in the current cognitive set (and an associated change in the underlying functional network configuration).

Bressler and Richter (2014) propose that cortical areas recruited under task-specific conditions may be linked by inter-areal beta synchrony to form NeuroCognitive Networks (self-organizing, large-scale distributed cortical networks) at the highest hierarchical levels. They propose that beta activity may serve the dual purpose of maintenance of such networks, and carrying top-down signals to lower levels of the cortical hierarchy. I suggest that the construction of a sentence-level meaning representation during unification entails the formation of such a NeuroCognitive Network (NCN), encompassing areas in left inferior frontal cortex, left temporal cortex, and left inferior parietal cortex (Hagoort, 2013, 2014), along with other relevant areas outside the core language network depending on the particular context in which sentence-level meaning construction is taking place (e.g., recruitment of the theory of mind network for taking another person's perspective). By 'sentence-level meaning' I am referring not just to the semantics associated with the individual words comprising a sentence, but also to the semantics derived from the syntactic structure governing the hierarchical relations between those words. Along the lines of the proposal of Engel and Fries (2010), a NCN constitutes the neural implementation of the current cognitive set (Bressler & Richter, 2014). Beta increases indicate that the current NCN configuration is being actively maintained, while beta decreases indicate that the current NCN configuration is under revision/change. For language comprehension, a decrease in beta power signals a change in the current NCN as a result of some cue in the linguistic input indicating to the system that the current sentence-level meaning representation needs to be revised.

1.2.3 Re-evaluating the evidence

Under this 'beta-maintenance' hypothesis, we can easily account for the cases of syntactic and semantic violations (and similarly for rhythmic 'violations' and the case of 'Unagreement' in

Spanish, where although not ungrammatical the agreement mismatch would constitute an unexpected event for the system) by realizing that they would act as cues to the language comprehension system indicating that the current sentence-level meaning under construction is incorrect in some way and needs to be revised. This would result in a change in the underlying NCN and hence lead to the beta power decrease observed relative to syntactically and semantically acceptable sentences. The cases where increased syntactic unification load (e.g., long distance subject-verb agreement dependencies) results in increased beta power can be dealt with if we accept that an increase in syntactic unification load may act as a cue to the language comprehension system indicating that the current NCN needs to be actively maintained. According to Engel and Fries (2010) this would result in an increase in beta power relative to the conditions with lower syntactic unification load (e.g., short-distance subject-verb agreement dependencies), and this is exactly what was observed.

Strong support for the 'beta-maintenance' hypothesis comes from a recent turn-taking experiment where participants listened to recordings of natural speech and had to press a button when they predicted that their interlocutor would finish their turn (Magyari, Bastiaansen, de Ruiter, & Levinson, 2014). Stimuli were constructed such that in one condition the turn-end was highly predictable, while in the other it was unpredictable. A large decrease in beta power (localized to left inferior frontal regions, so unlikely to be solely related to motor preparation) was present just before the key-press in the highly predictable condition, while in the unpredictable condition there was an increase in beta power before the key-press. In the highly predictable condition the language comprehension system predicts that the current NCN will soon need to change (in preparation for constructing a new sentence-level meaning representation) and that results in the decrease in beta power. In the unpredictable condition on the other hand, the language comprehension system is engaged in ongoing sentence-level meaning construction and has no reason to expect it to change yet, so the current NCN should be maintained and this results in the observed increase in beta power.

Notice that the 'beta-maintenance' hypothesis accommodates a large number of findings that cannot be accounted for by the beta-syntax mapping suggested by the ‘frequency-based segregation of syntactic and semantic unification hypothesis’ (henceforth, the 'beta-syntax' hypothesis). In fact, the beta-syntax mapping may be subsumed under the more domain-general 'beta-maintenance' hypothesis, such that certain (but not necessarily all) manipulations of syntactic processing may result in modulations of beta power related to the maintenance/change of the current NCN responsible for the sentence-level meaning under construction.

To reiterate, both the 'beta-syntax' hypothesis and the 'beta-maintenance' hypothesis propose that more demanding syntactic processing should result in higher beta power. The 'beta-maintenance' hypothesis claims that this is also the case when other types of processing (e.g., semantic processing) become more demanding. Similarly, both the 'beta-syntax' hypothesis and the 'beta-maintenance' hypothesis propose that when syntactic processing is disrupted, and it is clear to the system that the grammaticality of the sentence cannot be recovered, this should result in lower beta power. Again, the 'beta-maintenance' hypothesis claims that this is also the case for disruptions of other types of processing (e.g., semantic processing). Importantly, the 'beta-syntax' hypothesis proposes that when syntactic processing is temporarily disrupted (e.g., when there is ambiguity between alternative syntactic constructions and the linguistic input disambiguates towards the less preferred alternative) but the grammaticality of the sentence can still be recovered, beta power should be ***higher*** due to syntactic processing becoming more demanding after the disruption. Under the same circumstances, the 'beta-maintenance' hypothesis proposes that this disruption is taken as a cue that the representation of the underlying sentence-level meaning needs to change, and beta power should therefore ***decrease***.

This distinction rests on what I will refer to as the *strong* version of the 'beta-syntax' hypothesis, where beta oscillations are linked directly to syntactic structure building, which does not halt when the parser encounters input that disambiguates toward a less preferred syntactic construction at locally ambiguous regions of a sentence. An alternative possibility is what I will refer to as the *weak* version of the 'beta-syntax' hypothesis, where beta oscillations instead track disruptions of syntactic processing. In such cases, under this *weak* version of the hypothesis beta is expected to decrease upon encountering input that disambiguates toward a less preferred syntactic construction at locally ambiguous regions of a sentence. The *weak* version does not directly link beta oscillations to syntactic structure building, only to the detection of disruptions of syntactic structure building. This means that its predictions for how beta should be modulated when syntactic processing is disrupted are identical to those of the 'beta-maintenance' hypothesis. Importantly however, this *weak* version of the 'beta-syntax' hypothesis is not compatible with beta findings reviewed above for cases where syntactic processing load increases, because in such cases beta should not be modulated if it is indeed only related to detection of disrupted syntactic processing. In this thesis I will deal with the *strong* version of the 'beta-syntax' hypothesis unless explicitly stated otherwise.

1.3 Outline of the dissertation

This dissertation presents four empirical studies investigating various aspects of the relationship between beta oscillatory activity and language comprehension beyond the processing of individual words. These studies probe the hypothesized link between beta and syntactic processing, and one study directly compares this proposed role for beta with the more domain-general hypothesis that beta is related to maintenance/change of the current NCN responsible for the sentence-level meaning under construction.

Chapter 2 moves beyond isolated sentences to investigate how beta oscillatory activity is modulated by discourse-level information. In an EEG experiment Dutch short stories consisting either of 4 semantically related sentences forming a coherent discourse, or of 4 unrelated sentences, were employed. Manipulating discourse-level semantic coherence in this way addresses the question of how beta oscillatory activity, ostensibly related to sentence-level syntactic processing, is modulated by discourse-level information.

In *Chapter 3* the link between beta oscillatory activity and syntactic processing is explored further in a series of four EEG experiments comparing the processing of Dutch grammatical gender violations between native speakers of Dutch, and German late second language learners of Dutch. The first two experiments directly compare native speakers and late second language learners, while the following two experiments investigate how factors like composition of the stimulus set and task demands influence beta activity for the second language learners.

Chapter 4 takes a brief detour in an attempt to develop a reliable tool for detecting predictive lexical pre-activation during language comprehension (e.g., DeLong, Urbach, & Kutas, 2005; Szewczyk & Schriefers, 2013). Beta activity has been linked to top-down prediction in a predictive coding framework (Bastos et al., 2012; Friston, Bastos, Pinotsis, & Litvak, 2014), and such a tool would allow one to address questions of whether and when predictive pre-activation occurs, and thus to further probe beta in relation to predictive processing during language comprehension. This EEG study employed frequency-specific oscillatory entrainment to investigate whether or not memory reinstatement of so-called ‘frequency tags’ (Wimber, Maaß, Staudigl, Richardson-Klavehn, & Hanslmayr, 2012) associated with memory representations constitutes a reliable method for tracking lexical activation. Unfortunately, the method did not prove to be reliable enough to confidently apply it to investigations of predictive language processing.

In *Chapter 5* a direct comparison is made between the hypothesis that beta is directly related to syntactic processing and the more domain-general hypothesis that beta is related to the maintenance/change of the current NCN responsible for the representation of the sentence level meaning. A MEG study employed locally ambiguous Dutch subject- and object-relative clause sentences to investigate modulations of beta power at the disambiguating target word in object-relative clause sentences, for which the above hypotheses make opposing predictions.

Finally, *Chapter 6* summarizes the four experimental chapters, provides a discussion of the results of those chapters in light of the two hypotheses regarding the role of beta activity during language comprehension, and suggests how research in this area might proceed in the future.

1.3.1 A note on the structure of the dissertation

Each experimental chapter (*Chapters 2* to *5*) is written as a self-contained and independent journal article. Consequently, there is a certain amount of overlap, especially in the introductory text for each chapter. A combined bibliography containing references for all chapters directly follows *Chapter 6*. All tables and figures are numbered consecutively according to the chapter in which they appear (e.g., Figure 5.1 refers to the first figure in *Chapter 5*).

Chapter 2

Discourse-level Semantic Coherence Influences Beta Oscillatory Dynamics and the N400 During Sentence Comprehension

Based on:

Lewis, A.G., Schoffelen, J., Hoffmann, C., Bastiaansen, M. C. M., & Schriefers, H. (2016). Discourse-level semantic coherence influences beta oscillatory dynamics and the N400 during sentence comprehension. *Language, Cognition and Neuroscience*, Advance online publication.

Abstract

In this study we used electroencephalography to investigate the influence of discourse-level semantic coherence on electrophysiological signatures of local sentence-level processing. Participants read groups of four sentences that could either form coherent stories or were semantically unrelated. For semantically coherent discourses compared to incoherent ones, the N400 was smaller at sentences 2 to 4, while the visual N1 was larger at the third and fourth sentences. Oscillatory activity in the beta frequency range (13-21 Hz) was higher for coherent discourses. We relate the N400 effect to a disruption of local sentence-level semantic processing when sentences are unrelated. Our beta findings can be tentatively related to disruption of local sentence-level syntactic processing, but it cannot be fully ruled out that they are instead (or also) related to disrupted local sentence-level semantic processing. We conclude that manipulating discourse-level semantic coherence does have an effect on oscillatory power related to local sentence-level processing.

Keywords: language comprehension; discourse semantics; beta oscillations; N400

2.1 Introduction

Language comprehension is an inherently dynamic process, with multiple sources of linguistic and non-linguistic information impinging upon the interpretation of just about any utterance or text. Exactly when during comprehension these different sources of information play a role is still an open question. Until recently the majority of research on language comprehension has been aimed at understanding the processing of individual words or single sentences in isolation. This has especially been the case for the investigation of oscillatory neural dynamics related to language processing. Yet everyday language use typically takes place within far richer contexts, and the information being conveyed goes beyond the meaning that can be decoded from single words or isolated sentences. Here we take up the challenge, and investigate how discourse information affects oscillatory neural dynamics related to language comprehension.

Within the field of electroencephalography (EEG) research the analysis of event-related potentials/fields (ERPs/ERFs) has proven invaluable in exploring the timing of various types of linguistic processing (e.g., DeLong, Urbach, & Kutas, 2005; Friederici, 2002; Hagoort & van Berkum, 2007). However, ERPs/ERFs provide only a glimpse into the rich spatio-spectro-temporal dynamics contained in the EEG/MEG signal (Makeig, Debener, Onton, & Delorme, 2004). ERP/ERF analyses rely on averaging over trials and participants in order to improve the signal-to-noise ratio (SNR) to observe time- and phase-locked (to an event of interest) neural signatures in the EEG/MEG. However, not all neural activity related to an event is strongly phase locked (evoked activity) to that event, and measuring non-phase-locked (induced), oscillatory activity (Tallon-Baudry & Bertrand, 1999) in the EEG/MEG can provide additional or complementary information about the underlying cognitive processing.

Neural synchronization plays an important role in the coupling and uncoupling of functional brain networks (e.g., Pfurtscheller & Lopes da Silva, 1999; Singer, 1993; Varela, Lachaux, Rodriguez, & Martinerie, 2001). Functional networks are created by synchronous repetitive firing of populations of neurons, resulting in an increased probability that interacting neurons entrain one another in a rhythmic, frequency-specific manner (e.g., König & Schillen, 1991). This mechanism for the segregation of different types of information (represented in networks firing synchronously at different frequencies) also supports the integration (or binding) of information distributed over distant neural populations (Gray, König, Engel, & Singer, 1989). Measuring frequency-specific oscillatory neural dynamics provides us with a window onto the dynamic coupling and uncoupling of such functional networks, and how this changes depending on the cognitive task.

A number of studies have shown a link between oscillatory activity in the beta frequency range (13-30 Hz) and manipulations of syntactic processing (see Lewis, Wang, & Bastiaansen, 2015 for review). Bastiaansen & Hagoort (2006) reported higher beta power for centre-embedded relative clauses compared to syntactically less complex right-branching relative clauses. Similarly, Weiss et al. (2005) report higher beta coherence between anterior and posterior electrodes for object-relative clauses compared to syntactically less complex subject-relative clauses. Meyer, Obleser, and Friederici (2013) found higher beta power for long- compared to short-distance subject-verb agreement dependencies at the point in the sentence where the dependency could be resolved. They argued that this was related to syntactic unification (Hagoort, 2005, 2013), since syntactic unification is likely more difficult in the case of long-distance dependencies. Finally, a number of studies have shown that beta power is higher for syntactically legal sentences compared to sentences containing a syntactic violation at the target word (Bastiaansen, Magyari, & Hagoort, 2010; Davidson & Indefrey, 2007; Kielar, Meltzer, Moreno, Alain, & Bialystok, 2014; Kielar, Panamsky, Links, & Meltzer, 2014). Together, these studies show that when syntactic unification becomes more difficult beta power increases, while disrupting syntactic unification leads to a relative decrease in beta power.

There is now also a large body of evidence linking oscillatory activity in the gamma frequency range (the findings are somewhat variable in terms of the exact frequency range, but all fall within the classical 30-100 Hz gamma band) to semantic processing (see Lewis, Wang, & Bastiaansen, 2015 for review). Pena and Melloni (2012) for example, have shown that gamma power (55-75 Hz) increases while listening to sentences in one's own language, but not while listening to sentences in a language that one does not speak/understand (where semantic processing presumably does not take place). From a different perspective, van Berkum, Zwitserlood, Bastiaansen, Brown, and Hagoort (2004) reported higher gamma power for referentially correct words compared to words that had no referent or were referentially ambiguous (and hence the assignment of thematic roles was presumably disrupted). Finally, a number of studies have reported an increase in gamma power at a target word for words that can be meaningfully integrated with a strongly constraining prior sentence context, but no gamma increase either when the target word results in a semantic violation, or when the sentence context is not strongly constraining (Hald, Bastiaansen, & Hagoort, 2006; Penolazzi, Angrilli, & Job, 2009; Rommers, Dijkstra, & Bastiaansen, 2013; Wang, Zhu, & Bastiaansen, 2012; Weiss & Mueller, 2003 for coherence instead of power). Wang, Zhu et al. (2012) for instance compared sentences containing a high cloze probability (a measure of how well a particular target word completes a prior sentence context according to participants' offline judgements; Kutas & Hillyard, 1984) target word with the same sentences containing either a low cloze probability target word (weakly constraining sentence context) or a semantic violation. They showed a gamma power increase only in the high cloze condition. The findings discussed all show that gamma power increases whenever semantic unification is successful.

Some recent studies have begun to investigate how power in the beta and gamma frequency ranges evolves over the course of an unfolding sentence, and how this might change when sentence processing is disrupted. Bastiaansen, Magyari, and Hagoort (2010) compared syntactically legal sentences to sentences containing a syntactic violation (word category violation), and to randomized lists of the words contained in the legal sentences (little or no syntactic structure). They showed a linear increase in beta power across the sentence for the syntactically legal sentences and for sentences containing a syntactic violation, but only up to the point of the violation, at which time beta power started to return to baseline levels. There was no increase in beta power for the randomized word list condition containing no syntactic structure. Similarly, Bastiaansen and Hagoort (2015) compared syntactically and semantically legal sentences to randomized word lists (global syntactic violation) and to syntactic prose (syntactically legal sentences that are semantically uninterpretable because all content words are replaced by other unrelated content words; global semantic violation) within the same set of participants. They replicated the findings from Bastiaansen et al. (2010) for beta power (although this time there was a linear decrease across the sentence for the randomized word list condition and no linear trend for the syntactically legal sentences), and showed that gamma power was higher for semantically legal sentences compared to syntactic prose, but that there was no linear trend across the sentence in the case of the relationship between gamma power and semantic processing. Based on these findings Bastiaansen and Hagoort (2015) proposed the 'frequency-based segregation of syntactic and semantic unification' hypothesis, suggesting that synchrony in the beta and gamma frequency ranges might constitute separate channels for the simultaneous processing of syntactic and semantic information during language comprehension.

On the other hand, it is not clear that the 'frequency-based segregation of syntactic and semantic unification' hypothesis holds for all available data. In a similar paradigm, Wang, Jensen, et al. (2012) observed a beta (but not gamma) power difference when comparing target words in semantically anomalous sentences to those in semantically acceptable sentences. There are a number of other examples (see Lewis, Wang, & Bastiaansen, 2015 for discussion) of cases where a strict beta-syntax and gamma-semantics link does not appear to hold. While it therefore appears clear that there is a link between both beta and gamma oscillatory neural activity and sentence-level language comprehension, whether those links are exclusive to syntactic and semantic aspects of sentence processing respectively remains less clear.

Another electrophysiological signature of brain activity that is sensitive to semantic processing is the N400 event-related potential (ERP) component (see Kutas & Federmeier, 2011 for a recent review). The N400 is characterized by a negative-going deflection in the ERP waveform, typically peaking around 400 ms after the onset of a target word. An N400 can be observed in response to all content words in a sentence (along with other potentially meaningful stimuli), and the amplitude of the deflection is sensitive to a number of factors, most important of which for the present study is how easily a target word can be integrated into some preceding sentence context (cloze probability; Kutas & Federmeier, 2011). When comparing target words in a sentence with high cloze probability (good semantic fit; e.g. *The peanut was salted*) to words with low cloze probability (e.g., *The peanut was small),* or to semantically incongruous words (e.g., *The peanut was in love*), the amplitude of the N400 is reduced and the difference between conditions is termed an N400 effect, exhibiting a characteristic centro-parietal scalp distribution.

More recently it has been shown that discourse-level information can have an influence on local semantic processing within a sentence (see van Berkum, 2012) for review). For instance, Nieuwland and Van Berkum (2006) showed that by inserting a sentence containing a semantic animacy violation into a discourse context that changes reader's/listener's expectations about the animacy of the discourse referent (e.g., in the sentence from the previous paragraph *The peanut* is described in the preceding discourse as having animate characteristics), the direction of the N400 effect can be reversed (the N400 is now more negative when *salted* is the target word compared to *in love*). This is evidence that discourse-level information can have an effect on electrophysiological signatures related to the processing of semantics within a sentence.

Something that is not yet clear is exactly which discourse-level factors (only animacy or possibly other factors like discourse coherence or anaphora) can influence ERP signatures related to sentence-level processing, and under which circumstances. Another outstanding question concerns whether or not discourse-level factors can influence the oscillatory signatures that have been found for sentence-level semantic and syntactic unification. In the present study we aimed to address some of these questions by revisiting an existing dataset where participants read semantically coherent (COH) and incoherent (INCOH) short stories while their EEG was measured (Lewis, 2012). COH stories consisted of 4 sentences that fit together to describe a situation or event, and INCOH stories consisted of 4 unrelated sentences (see Table 2.1 for example stories).

Table 2.1 Example materials and their English translation (in italics).

Condition	*Example Materials*
COH	Charles verliet zijn vaderland Senegal om in Europa te werken.
	Charles left his home country Senegal to work in Europe.
	Met een levensgevaarlijk klein bootje werd hij naar Tenerife gesmokkeld.
	With a dangerously small boat he was smuggled to Tenerife.
	Hij moest daar hard werken voor een klein beetje geld.
	There he had to work hard for very little money.
	Zijn familie had het geld dat hij stuurde hard nodig.
	His family desperately needed the money that he was sending.
INCOH	Charles verliet zijn vaderland Senegal om in Europa te werken.
	Charles left his home country Senegal to work in Europe.
	Een avond hadden ze een taart achtergelaten in de keuken.
	One evening they left a hot pie in the kitchen.
	Toevallig kwam een agent de hoek om die hen arresteerde.
	Coincidentally a cop came around the corner that arrested them.
	Maar na een jaar moest hij al naar de sloop.
	But after just a year it was ready for the dump.

Notes: COH: semantically coherent condition; INCOH: semantically (discourse-level) incoherent condition; *italics*: English translation.

Our original approach to the analysis, investigating the evolution of power over the course of the entire story, did not yield any statistically significant results (Lewis, 2012; see also Lewis et al., 2015 for discussion). We have argued that one reason for this may be the relatively poor SNR in the data due to the unusually long trials (23200 ms for the word by word presentation of the whole story). In the present study, we improve the SNR by taking the average over all words within each sentence and each condition, thus retaining temporal information in the form of the sentence number (first, second, third, or fourth sentence of the story), but averaging out as much noise as possible. This approach has the added benefit that it addresses the within-sentence temporal variability of a potential effect of our experimental manipulation. Although we can be certain that discourse coherence breaks down at sentences 2, 3 and 4 for the INCOH condition (see *Section 2.2.2*, rating task), it turned out to be impossible to tightly control the exact point within each sentence at which this occurred (e.g., coherence might break down at word number 5 of the fourth sentence for one stimulus item but at word number 7 of the fourth sentence for another stimulus item, and the point of coherence break down for a given stimulus will presumably even vary between participants). By averaging over all words within each sentence we effectively remove this issue, at the cost of a potential loss of sensitivity due to the inclusion of words where an effect of our manipulation does not occur. This likely makes our statistical analyses particularly insensitive to the detection of potential interaction effects.

We performed a time-frequency analysis of power changes relative to a pre-story baseline period in partially overlapping low (2-30 Hz) and high (28-100 Hz) frequency ranges. This allows us to test whether discourse-level semantic coherence has any effect on local sentence-level processing, reflected in differences in beta and/or gamma power. We also performed an ERP analysis as we suspected that the N400 might be sensitive to our semantic coherence manipulation. The ERP analysis was also performed with the data averaged over all words within each sentence and each condition.

We had two main hypotheses for this experiment. First, we hypothesized that the N400 ERP component should be sensitive to our semantic manipulation, and thus should be larger for the INCOH than the COH condition at sentences 2, 3, and 4. This will serve as an indication of whether or not our manipulation of discourse semantics has any effect on online semantic processing. Second, we expect to find higher gamma and/or beta power in the COH than in the INCOH condition for sentences 2, 3, and 4. Gamma power modulations related to our experimental manipulation would clearly reflect disrupted local semantic processing. Modulations of beta power on the other hand are less clear, because as we have outlined above, beta has been observed for manipulations of both syntactic and semantic processing.

2.2 Methods

2.2.1 Participants

Thirty native speakers of Dutch took part in the experiment, 20 of whom were included in the final analysis (7 males, 13 females; aged 18 to 27). Participants provided informed consent and were paid or equivalently rewarded with course credits for their participation. All participants reported normal or corrected-to-normal vision, and were right handed. None of the participants reported any neurological impairment, nor had they participated in any of the previous experiments involving the construction of the stimulus materials.

Five participants were excluded from the final analysis due to recording problems. Five other participants were excluded due to poor data quality (more than 37.5 % of trials rejected in either condition). The relatively high number of excluded participants is primarily due to the fact that the experimental trials had a considerably longer duration than in experiments on isolated sentences, thus leading to more eye movements and other movement artefacts.

2.2.2 Stimulus materials

All stimuli consisted of Dutch short stories, each comprised of four syntactically and semantically acceptable sentences. Every sentence contained exactly ten words (Table 2.1). Conditions differed in terms of whether the sentences comprising the stories formed a coherent discourse (COH), or were unrelated to one another (INCOH). Two additional conditions where the discourse became incoherent starting at the third and fourth sentences respectively were included as fillers.

Eighty coherent and 80 incoherent stories (specifications just described) were taken from Hoffmann (2011). For the INCOH condition, the second, third, and fourth sentences in the COH stories were randomly exchanged across items (i.e., sentence 2 from the first story was exchanged with sentence 2 from one of the other 79 COH stories; sentence 3 from the first story was exchanged with sentence 3 from one of the other 79 COH stories; etc.). Latent Semantic Analysis (LSA) scores (Landauer, Foltz, & Laham, 1998) were calculated to confirm that this led to low semantic coherence between the sentences in this condition (Hoffmann, 2011). These 80 COH and 80 INCOH stories were counterbalanced across participants using two experimental lists, each comprised of 40 COH and 40 INCOH stories. This meant that no participant read the same sentence more than once in the experiment, while all sentences appeared an equal number of times at the same sentence position in both conditions. Importantly, this shuffling procedure was applied at the level of sentences, so that word position within a sentence and overall sentence structure was preserved across participants. The only factor that was manipulated for each of sentences 2 to 4 was whether or not that sentence fit coherently with the preceding sentence (or sentences). For the fillers, the third and fourth (10 stories), and fourth (10 stories) sentences respectively of an additional 20 coherent stories not used in the main comparisons were randomly exchanged to induce a breakdown in semantic coherence starting at sentence 3 or at sentence 4.

Table 2.2 Results from the rating task.

	Sentence 2		*Sentence 3*		*Sentence 4*	
Condition	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
COH	6.6	0.4	6.6	0.3	6.7	0.3
INCOH	2.3	0.5	1.7	0.5	1.4	0.3
INCOH 3	6.4	0.6	1.6	0.6	1.3	0.4
INCOH 4	6.3	0.5	6.3	0.6	1.6	0.5

Notes: COH: semantically coherent condition; INCOH: semantically (discourse-level) incoherent condition; INCOH 3 and INCOH 4: fillers; rating of 7: fits perfectly with previous sentence; rating of 1: does not fit with previous sentence at all.

A rating task was performed with a group of participants who did not take part in the EEG experiment (Hoffmann, 2011). Their task was to rate how well each sentence comprising the stories fit with the previously presented sentence on a seven point Likert scale, with 1 being "not at all", and 7 being "very well". The results[1] are shown in Table 2.2 and indicate that for the INCOH and two filler conditions participants were already aware of the semantic coherence breakdown at the end of the second, and third or fourth sentences of the stories respectively.

2.2.3 Experimental design and procedure

Participants were tested in a dimly lit, sound-attenuating and electrically shielded booth. They were seated comfortably in front of an LCD computer monitor (Samsung SyncMaster 940eW), with a viewing distance of between 70 and 80 cm. Letters were presented in white on a black background using a 20-point sized Consolas font type. All words subtended a visual angle of 3.13° vertically.

Sentences were presented word by word in the centre of the screen. For each sentence, the first letter of the first word was capitalized, and the final word was presented with a period. A single trial consisted of an entire story containing four sentences, a movement cue, and a fixation cross. Words were presented for 300 ms, followed by a 200 ms blank screen between words. Each trial began with the presentation in the centre of the screen of three asterisks separated by two spaces for 4000 ms, indicating that participants could move their eyes and blink. This was immediately followed by a fixation cross presented in the centre of the screen for 3000 ms, indicating that eye movements and blinking should be avoided, and that the story was about to begin. The first word of the first sentence immediately followed the fixation cross. Each sentence lasted 5000 ms and was followed by an 800 ms inter-sentence blank screen before the onset of the first word of the next sentence. For the last sentence of each story, the inter-sentence interval was immediately followed by a new trial. A single trial lasted 30200 ms (including fixation and blinking periods).

Participants were instructed to read all stories attentively for comprehension, and to continue reading regardless of whether or not the story made sense to them. They read a total of 100 stories (40 COH, 40 INCOH, and 20 fillers), presented in 20 blocks of five stories each. For 50 percent of the participants, stories in an experimental list were presented in reverse order to control for potential order effects. Ten training stories were presented to participants before the experiment.

2.2.4 EEG recordings

Participants were fitted with a 64 electrode actiCap with electrodes positioned according to the standard 10/20 system. EEG signals were recorded using 60 Ag/AgCl active sensors mounted in the cap and referred to the right mastoid. An additional electrode was placed on participants' left mastoid for re-referencing offline, and a ground electrode was placed on the centre of the forehead. An additional electrode was placed on the suborbital ridge of participants' left eye for recording eye-blinks.

Electrode impedance was kept below 10 kΩ. EEG and EOG recordings were amplified using BrainAmp DC amplifiers (Brain Products GmbH, Gilching, Germany) with a band-pass filter of 0.053 to 249 Hz, digitized online with a sampling frequency of 1000 Hz and stored for offline analysis.

2.2.5 Data pre-processing

EEG data were analysed using the FieldTrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011) running in a MatLab environment (R2012a; Mathworks, Inc.). For each participant, scalp electrodes were re-referenced to the average of electrodes placed on the left and right mastoid (linked-mastoid reference). A band-stop filter was applied at 50, 100, and 150 Hz in order to minimize the effects of power line interference (50 Hz) and data were segmented from -2500 to 24000 ms relative to the onset of each story.

Next the data were decomposed into independent components (ICA using the 'runica' implementation in FieldTrip with default settings), resulting in 60 component time courses. Components which captured eye-blinks and eye movements were removed, and the remaining components were recombined (Jung et al., 2000; Makeig, Jung, Bell, Ghahremani, & Sejnowski, 1997). Between 0 and 2 components were removed per participant.

Trials still containing muscle artefacts were rejected by visual inspection of the data, band-pass filtered between 110 and 140 Hz (this frequency range is where the majority of the energy associated with muscle activity is concentrated). Any linear trends in the data were removed and a baseline correction was applied to every trial using an interval from -2500 to 0 ms relative to story onset. Trials with amplitude higher than 75 µV or lower than -75 µV were excluded from further analysis. There was no significant difference between the number of trials in the two conditions (COH: M = 29.4, SD = 2.72; INCOH: M = 29.05, SD = 3.24; p = .66). Finally, data were segmented from -1000 to 1000 ms relative to the onset of every word of each sentence comprising the stories for the COH and INCOH conditions combined.

2.2.6 Event-related potential (ERP) analysis

Data for each word were high-pass filtered above 0.1 Hz and low-pass filtered below 30 Hz using a windowed sinc finite-impulse response filter with FieldTrip default settings. Next, data were segmented into COH and INCOH conditions, and within each condition into sentences 1 to 4 respectively. A condition-specific baseline correction using a period from -200 to 0 ms relative to word onset was applied to the data for each individual word. Finally, data were averaged within each of the four sentences for the two conditions separately from -100 to 500 ms relative to word onset to obtain participant-specific ERP waveforms.

2.2.7 Time-frequency and inter-trial coherence analyses

A multitaper approach (Mitra & Pesaran, 1999) was used to compute time-frequency (TF) representations for the single trial data of each participant. TF representations were calculated in two partially overlapping frequency ranges (Womelsdorf, Fries, Mitra, & Desimone, 2006) because of the trade-off between time and frequency resolution that results from this multitaper approach.

In a high-frequency range (28-100 Hz), 250 ms time-smoothing and 8 Hz frequency-smoothing windows applying a Slepian taper sequence were used to calculate power changes in frequency steps of 4 Hz and time steps of 10 ms. Each time point in the resultant TFRs is thus a weighted average of the time points ranging from 125 ms before to 125 ms after this time point. In a low-frequency range (2-30 Hz), 400 ms time-smoothing and 2.5 Hz frequency-smoothing windows using a Hanning taper were applied in frequency steps of 1 Hz and time steps of 10 ms. Single-trial Fourier spectra were averaged for each participant from 0 to 500 ms relative to the onset of each word of the stories. Data epoch length is limited by the stimulus onset asynchrony (SOA) of 500 ms in order to avoid averaging overlapping data segments. This resulted in a TF representation of power for each participant averaged over all words comprising the stories, irrespective of condition or sentence number. These participant averages were then expressed as a relative change (in dB) from a baseline period between 750 and 250 ms prior to story-onset (fixation). The average TF representation of power over all participants and scalp electrodes was then calculated for visual inspection.

For the low frequency range we also computed the inter-trial coherence (ITC; Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996) for each participant from 0 to 500 ms relative to the onset of each word of the stories, by first normalizing the Fourier spectrum of each trial by its amplitude and then averaging the result across all trials for that participant. This provides a frequency-resolved measure of the degree of trial-to-trial phase consistency over time (Makeig et al., 2004). We used this to distinguish evoked activity (strongly phase-locked and likely related to the ERP results) from induced activity (time- but not phase-locked) in subsequent TF analyses.

Resultant participant-specific ITC values were then averaged over all participants and scalp electrodes for visual inspection.

Next we selected TF ranges of interest for the low frequency range based on previous literature and on visual inspection of the TF and ITC data averaged over all words, all participants, and all scalp electrodes (Figure 2A). In this way we were able to select regions of interest without statistical comparison between conditions at this stage (there is no condition-specific information in this representation of the data), thus avoiding double-dipping later in the statistical analysis. Our criteria were: 1) a visible increase or decrease in power in the TF representation relative to baseline (to establish the presence of a power modulation compared to baseline); 2) only weak (less than 0.15) or no phase-locking visible for the corresponding TF range in the ITC values (to exclude the potential contribution of ERP components to the TF representations of the induced oscillatory activity of interest); 3) good correspondence with previous results in terms of frequency range selected (to ensure any potential effects make sense in light of previous literature). This resulted in the selection of alpha/theta (6-12 Hz; 350-500 ms relative to word onset), alpha/beta (9-17 Hz; 0-140 ms relative to word onset), early beta (21-28 Hz; 40-160 ms relative to word onset), and late beta (13-21 Hz; 260-480 ms relative to word onset) TF ranges of interest for further analysis (see black boxes in Figure 2.2A). In addition, we selected a TF range of interest for the large theta power increase (4-7 Hz; 150-300 ms) despite the ITC data clearly indicating that this is phase-locked activity, in order to investigate whether or not this is related to potential ERP findings (for discussion see Bastiaansen, Mazaheri, & Jensen, 2012).

For the high frequency range we performed statistical analyses on the entire range (28-100 Hz), as well as for the mean power in a low gamma frequency range (35-55 Hz) based on the majority of findings relating gamma power to semantic processing (Lewis et al., 2015).

Single-trial Fourier spectra per participant were then segmented into COH and INCOH conditions, and within each condition into sentences 1 to 4, from 0 to 500 ms relative to word onset. Fourier spectra were averaged, resulting in participant-specific averages for sentences 1 to 4 for the COH and INCOH conditions respectively. These participant averages were then expressed as a relative change (in dB) from the baseline period between 750 and 250 ms prior to story-onset (fixation). This provides us with a measure of the average relative power change from baseline for all words within each sentence of the stories for the COH and INCOH conditions separately. By averaging over all words in a sentence we improve the SNR and at the same time take into account temporal variability in the potential effect of our coherence manipulation on theta, alpha, and beta power (e.g., appearing at word 7 in one trial and word 8 in another trial).

2.2.8 Statistical analyses

The statistical significance of all comparisons was evaluated using a cluster-based random permutation approach (Maris & Oostenveld, 2007). We used this approach because of its natural handling of the multiple comparisons problem (MCP).

Cluster-based random permutation statistics control the family-wise error rate by making use of the spatial, spectral, and temporal autocorrelation in EEG data. In short, a dependent-samples T-test is performed for every data point (electrode-frequency-time point for TF or electrode-time point for ERP data) giving uncorrected P-values. A pre-set significance level is chosen (here 5% single-tailed for the N400 ERP analysis; 5% two-tailed, for all other comparisons) and any data points not exceeding this level are discarded (set to zero). Clusters are calculated from the remaining data points based on their adjacency in space (adjacent electrodes), time, and frequency. Cluster-level statistics are then calculated by summing the values of the T-statistics for all data points in each cluster. A permutation distribution is created by randomly assigning participant averages to one of the two conditions 3000 times, and each time calculating cluster-level statistics as just described. The highest cluster-level statistic from each randomization is entered into the permutation distribution and the cluster-level statistics calculated for the measured data are compared against this distribution. Clusters falling in the highest or lowest 2.5^{th} percentile of the estimated distribution were considered significant (lowest 5^{th} percentile for the N400 ERP analysis).

We compared the COH and INCOH conditions separately for each sentence comprising the stories. For the ERP data we hypothesized that our manipulation of discourse semantics should result in a more negative N400 peak in the INCOH than in the COH condition for sentences 2 to 4. Our statistical comparison was therefore based on the mean ERP amplitude in a time window (300 to 500 ms relative to word onset) typically capturing N400 effects (see e.g., Kutas & Federmeier, 2011). To test whether there were any earlier ERP effects we tested the time window between 0 and 300 ms relative to word onset, now forming clusters in time as well as space. For the TF data we compared mean power values in the selected theta (4-7 Hz; 150-300 ms relative to word onset), alpha/theta (6-12 Hz; 350-500 ms relative to word onset), alpha/beta (9-17 Hz; 0-140 ms relative to word onset), early beta (21-28 Hz; 40-160 ms relative to word onset), and late beta (13-21 Hz; 260-480 ms relative to word onset) TF ranges, forming clusters only in space. We also compared COH and INCOH conditions for the entire gamma frequency range (28-100 Hz), clustering in space, frequency, and time, as well as in a low gamma frequency range of interest (35-55 Hz), clustering in time and space.

For observed statistically significant differences between COH and INCOH conditions in both the TF and ERP analyses we tested for an interaction between condition (COH/INCOH) and sentence position (Sentence1/Sentence2/Sentence 3/Sentence4) by extracting mean power or amplitude values respectively in the TF region or time window of interest, averaged over all electrodes identified based on the output of the cluster-based statistics from the sentence exhibiting the largest effect. These values were entered into a repeated measures ANOVA with condition and sentence position as factors. There were no cases where sphericity was violated according to Mauchley's test. We only interpreted interaction effects, but not main effects, in order to avoid double-dipping. Interactions were not broken down further since we already have pairwise comparisons from the cluster-based statistical output.

2.3 Results

Statistical comparisons were made between COH and INCOH conditions separately for each of the four sentences comprising the stories. Interactions between sentence position and condition were also tested. We expected our statistical analyses to be relatively insensitive to potential interaction effects, due to the poor SNR and averaging over all words within a sentence (as already discussed, effects of manipulating discourse-level semantic coherence are not likely to be present at all words of the sentences). We therefore still describe differences between conditions at each sentence position, but are careful not to make inferential claims about how this differs from sentence to sentence in cases where the interaction is not significant.

2.3.1 Event-related potential results

ERP effects were quantified by differences in mean amplitude in the N400 time window (300 to 500 ms relative to word onset), or by temporally and spatially contiguous time points identified by the cluster-based permutation approach in an earlier time window (0 to 300 ms relative to word onset). Figure 2.1 shows the ERP waveforms for the COH and INCOH conditions for each of the four sentences comprising the stories (left column), along with the scalp distribution of the difference between conditions (INCOH minus COH) for the early and N400 time windows (middle and right column respectively).

2.3.1.1 N400 time window

In the N400 time window we observed a statistically significant negative difference between the INCOH and COH conditions at the third ($p = 0.03$), and fourth ($p < 0.001$) sentences, with this difference exhibiting a trend toward significance at the second sentence ($p = 0.05$). The interaction between condition and sentence position was statistically significant ($p = 0.016$). This, in combination with the cluster-based pairwise comparisons, suggests that sentences 2-4 exhibit differences between COH and INCOH conditions, while the first sentence does not. The INCOH condition exhibits a larger negative-going deflection compared to the COH condition between about 350 and 450 ms after word onset, with a centro-parietal scalp distribution for the difference (Figure 2.1). Based on the timing and scalp distribution we identify this as an N400 effect.

2.3.1.2 Early time window

In the early time window we observed a significant positive difference between the INCOH and COH conditions at the third ($p = 0.02$) and fourth ($p = 0.02$) sentences. The interaction between condition and sentence position was not statistically significant ($p = 0.75$). Based on these results we cannot make any inferential claims about whether the difference between COH and INCOH conditions differs across sentence positions. Any discussion of such effects for the ERP data in the early time window is thus purely descriptive in nature. The COH condition shows a larger negative-going deflection compared to the INCOH condition over a wide range of electrodes (Figure 2.1). This effect appears to begin around 80 ms relative to word onset and lasts until about 200 ms. Closer examination reveals that this difference is already present at the second sentence, but that there it is smaller and less widely distributed across the scalp. As a result, it is not significant there ($p = 0.22$). Based on the timing and scalp distribution of this difference we argue that it is likely a visual N1 effect.

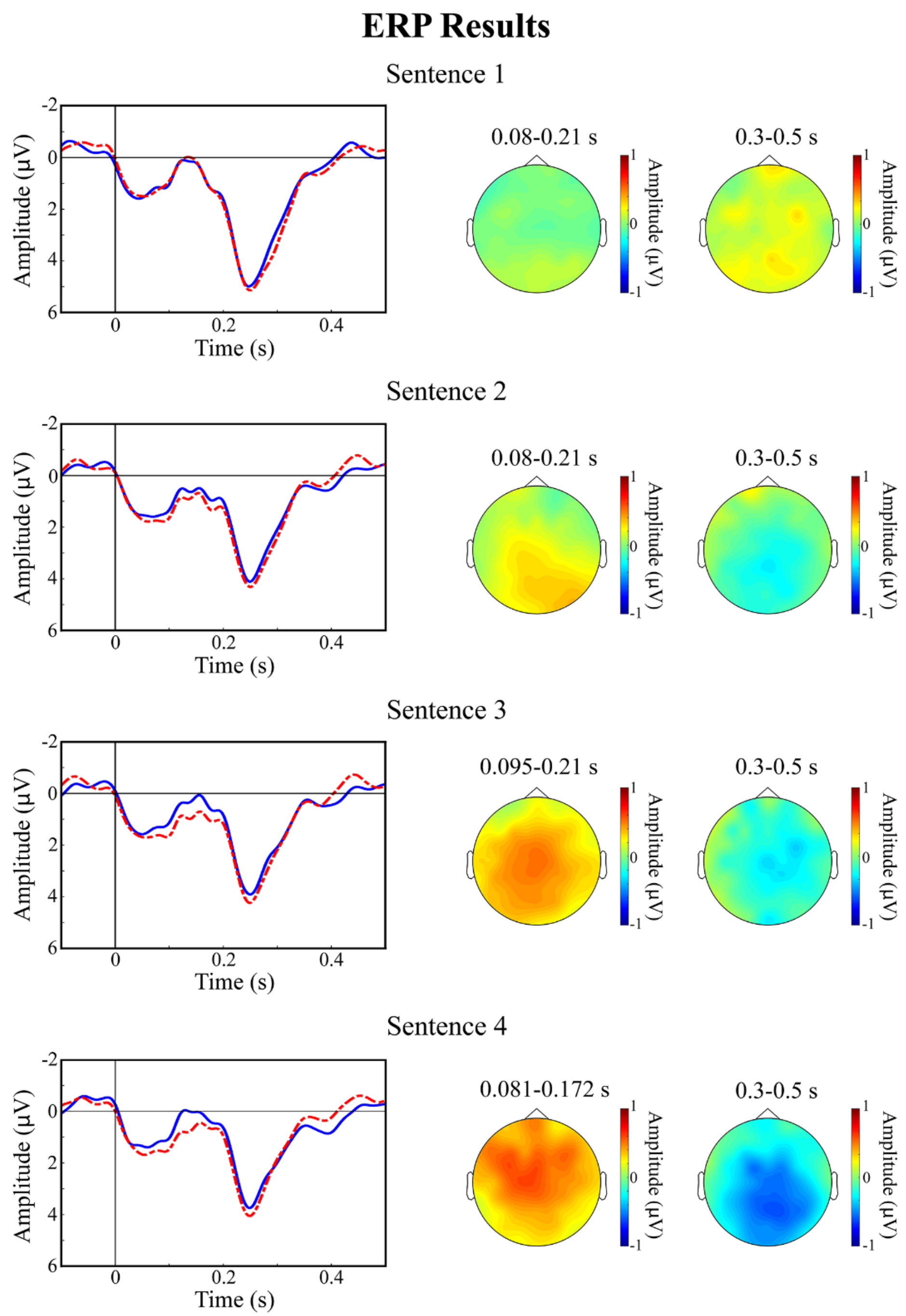

Figure 2.1 Results of the ERP analysis. The results for the average over all words in each sentence are presented separately. The left column shows ERP time courses for a representative electrode CPz (blue solid = COH; red dashed = INCOH). The middle column shows the scalp distribution of the difference between conditions (INCOH minus COH) averaged over a time window corresponding to the significant effects (sentences 3 and 4) found in the early time window analysis. The third column shows the scalp distribution of the difference between conditions (INCOH minus COH), now averaged over the N400 time window (300 ms to 500 ms relative to word onset), exhibiting a statistically significant effect at sentences 3 and 4, and a trend at sentence 2.

2.3.2 Time-Frequency results

For the low frequency range, we selected five TF ranges of interest for statistical comparison, a theta (4-7 Hz; 150-300 ms relative to word onset), an alpha/theta (6-12 Hz; 350-500 ms relative to word onset), an alpha/beta (9-17 Hz; 0-140 ms relative to word onset), an early beta (21-28 Hz, 40-160 ms relative to word onset), and a late beta (13-21 Hz; 260-480 ms relative to word onset) range. Figure 2.2A shows the TF representation of power (top) and corresponding ITC values (bottom) averaged over all words in the stories irrespective of condition or sentence. TF ranges of interest are marked by black boxes.

There were no statistically significant differences between the COH and INCOH conditions for any of the 4 sentences in the theta, the alpha/theta, the alpha/beta, or the early beta TF ranges of interest. In the late beta TF range there was a significant difference between COH and INCOH conditions at the fourth sentence of the stories ($p = 0.02$). The interaction between condition and sentence position was not statistically significant ($p = 0.86$). Based on these results we cannot make any inferential claims about whether the difference between COH and INCOH conditions differs across sentence positions. Any discussion of such effects for the TF data in the late beta frequency range is thus purely descriptive in nature. Figure 2.2B shows the scalp distribution of the difference between conditions (COH minus INCOH) for each of the four sentences in the late beta TF range, along with bar plots of the power decrease for each condition at sentences 1 to 4, averaged over all electrodes contributing to the statistically significant difference at the fourth sentence. Sentence 4 clearly exhibits the strongest and most widespread difference, with maxima over left fronto-central, right frontal, and right temporal electrodes. Sentence 3 also exhibits a difference, but it only shows a trend towards statistical significance ($p = 0.06$), with maxima over left fronto-central and right tempro-parietal electrodes. The effect is driven by a decrease in beta power relative to baseline in both conditions (Figure 2.2A and bar plots in Figure 2.2B), which is stronger in the INCOH than in the COH condition.

For the high frequency range, we selected the entire frequency range (28-100 Hz) as well as a low gamma frequency range of interest (35-55 Hz) for statistical comparison. There were no statistically significant differences between the COH and INCOH conditions for either of these ranges in any sentence position. This suggests that local sentence-level gamma power was not sensitive to our manipulation of discourse-level semantic coherence.

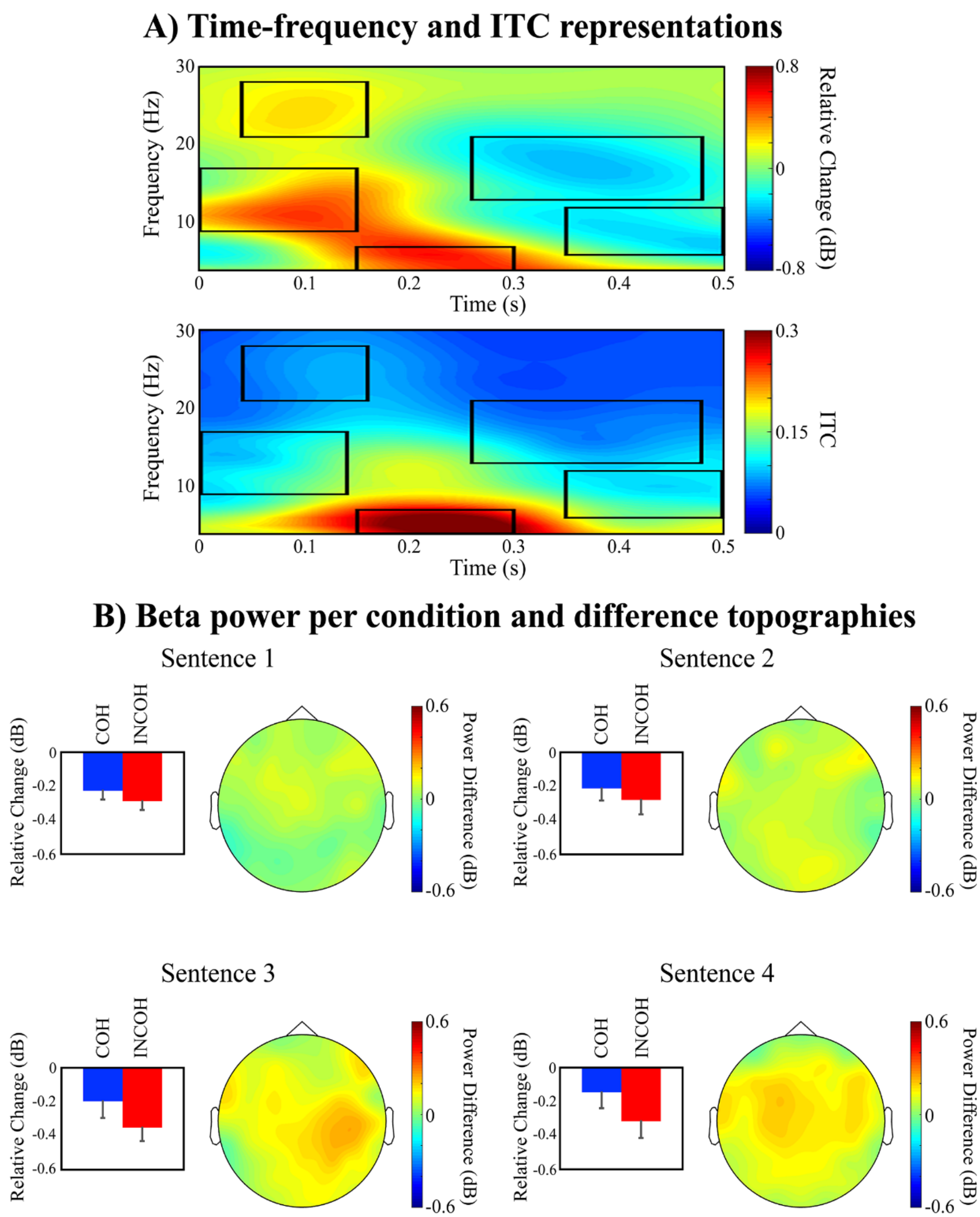

Figure 2.2 Results of the TF analysis of power. (A) TF (top) and ITC (bottom) representations for the average over all participants, all scalp electrodes, and all words in all sentences in both the COH and INCOH conditions. Black boxes indicate the TF ranges of interest selected for statistical testing. (B) Bar plots of mean power in the beta TF range of interest (13-21 Hz; 260-480 ms relative to word onset) for each condition averaged over electrodes contributing to the statistically significant difference at sentence 4, as well as scalp distributions of the difference between conditions (COH minus INCOH) averaged over all words within each of the four sentences for the beta TF range of interest. This difference is statistically significant at sentence 4 and shows a trend towards significance at sentence 3. Error bars on the bar plots indicate standard error of the mean.

2.3.3 Correlation analysis

In order to test whether there is any direct relationship between our beta oscillatory effects and our N400 ERP findings, we performed a Pearson correlation between N400 amplitude differences and beta power differences between the COH and INCOH conditions for sentences 2 to 4. For every participant, and for each sentence separately, we extracted mean power difference values from the late beta TF range of interest, averaged over all electrodes contributing to the statistically significant difference at the fourth sentence. We also extracted mean amplitude differences per participant for each sentence, averaged over the N400 time interval and over all electrodes contributing to the statistically significant difference at the fourth sentence. A Pearson correlation was then performed across participants between these two values. This correlation analysis did not produce any statistically significant correlations between beta power and N400 amplitude differences. We can thus conclude that there is no direct relationship between our beta oscillatory and N400 effects related to our experimental manipulation of discourse-level semantic coherence.

2.4 Discussion

The frequency-based segregation of syntactic and semantic unification hypothesis (Bastiaansen & Hagoort, 2015) claims that there is a close relationship between syntactic processing and oscillatory neural dynamics in the beta frequency range, and between semantic processing and oscillatory activity in the gamma frequency range. On the other hand, modulations of beta power have also been observed for manipulations of semantic processing (e.g., Wang, Jensen, et al., 2012). Furthermore, discourse-level information can influence electrophysiological brain signatures related to local sentence-level semantic processing (i.e. the N400 ERP component; Nieuwland & Van Berkum, 2006). We tested whether a different discourse-level factor, semantic coherence, can influence local sentence-level processing as indexed by modulations of the N400, as well as beta and gamma oscillatory power. Participants read groups of 4 sentences that either formed coherent stories (COH) or were semantically unrelated (INCOH), as indicated by both LSA scores (Hoffmann, 2011) and a rating task (Table 2.2).

The ERP analysis produced an N400 effect at sentences 2 to 4, with more negative-going waveforms in the INCOH compared to the COH condition, and a visual N1 effect that was only significant at sentences 3 and 4, with a more negative peak for the COH compared to the INCOH condition (Figure 2.1). The time-frequency analysis of power produced a single result in the beta frequency range (13-21 Hz), with higher beta power in the COH compared to the INCOH condition at sentence 4 between 260 and 480 ms relative to word onset, exhibiting a frontal, fronto-central, central, and centro-parietal scalp distribution. At sentence 3 a trend towards significance was present, exhibiting a similar scalp distribution. There were no gamma power effects related to our manipulation of sentence-level semantic coherence. Importantly, we observed an interaction between sentence position and condition only for the N400 effect, and thus any discussion of differences between the COH and INCOH conditions differing across sentence position for the N1 ERP and for the beta TF findings is purely descriptive in nature.

2.4.1 Discourse-level semantic coherence influences the N400

The main purpose of the ERP analysis was to confirm that our discourse-level semantic coherence manipulation had an effect on online sentence-level semantic processing. We deliberately avoid entering discussions about the nature of the processing giving rise to N400 effects (e.g., lexical retrieval difficulties, disruption of semantic integration, or lower predictability; see Kutas & Federmeier, 2011), as we do not think our experimental manipulation allows us to add anything to this debate.

As hypothesized, the N400 was sensitive to our discourse-level semantic manipulation as soon as the stories became incoherent (sentences 2 to 4). It is clear from the scalp distribution in Figure 2.1 (right-most column) that the effect becomes larger and more widespread across the scalp (this does not necessarily indicate that more brain regions become involved, but can also result from increased activity at the same underlying sources) as we move from sentence 2 to 4. The N400 is classically related to semantic processing at the level of single words, as well as the sentence level (and also certain non-linguistic stimuli; Kutas & Federmeier, 2011), and has also been shown to be sensitive to discourse-level semantic information in the form of the animacy assigned to a particular referent based on the preceding discourse (Nieuwland & Van Berkum, 2006). Here we show that a different discourse-level factor, semantic coherence between sentences comprising a short story, can also have an influence on the N400. We argue that this is related to local sentence-level semantic processing. Bearing in mind that we average over all words within each sentence for the two conditions separately, it is likely that our manipulation of semantic coherence has an effect on the N400 at one, or more likely at a few words within each sentence (probably a variable number of words in each sentence across items, and perhaps even participants), and that this is what drives the effect when averaging over all words for sentences 2 to 4. This would suggest that manipulating semantic coherence at the discourse level disrupts local sentence-level semantic unification, possibly related to difficulties with processing anaphoric relations from sentence to sentence (van Berkum, Koornneef, Otten, & Nieuwland, 2007) and/or with thematic role assignment (Paczynski & Kuperberg, 2011). A question that should be addressed in future work is which of these aspects of sentence-level processing is affected by manipulating discourse-level semantic coherence. For instance, more fine-grained stimuli could be used to specifically target the breakdown of anaphoric relations between sentences (e.g., Koornneef & Sanders, 2013).

St. George, Mannes, and Hoffinan (1994) report a similar finding for ambiguous paragraphs preceded by a disambiguating title compared to the same paragraphs without a title. The presence of a title causes the sentences comprising the paragraphs to be more coherent, and disambiguates the paragraph for the reader. They report an N400 effect, with a more negative-going N400 component for the condition without a title (ambiguously related sentences) compared to the condition with a title (coherent sentences), exactly in line with our N400 effect. Importantly, they also report an enhanced P1-N1 ERP component for their condition with a title. They interpret this finding in terms of a better ability to allocate attention to the condition where sentences are more coherent due to the presence of a title. This directly corroborates our visual N1 effect at sentences 3 and 4 (Figure 2.1). We observed a larger N1 in the COH compared to the INCOH condition between about 80 and 200 ms relative to word onset. Since this ERP component has been linked to the allocation of visual attention (e.g., Vogel & Luck, 2000), we conclude that our participants allocate more attention to the words in each sentence in the COH condition compared to the INCOH condition for sentences 3 and 4. This is likely because that information is more relevant in the COH condition, when a detailed situation model has to be constructed, whereas in the INCOH condition sentences can simply be read without any further processing related to construction of a detailed situation model. This becomes clearer the later in the stories a sentence appears, which might be the reason the N1 difference appears to become larger at later sentences (Figure 2.1).

2.4.2 Higher beta power for coherent short stories

Our manipulation of discourse-level semantic coherence had an effect on beta power at the third and fourth sentences of the stories (Figure 2.2B), but only reached statistical significance at the fourth sentence. The effect at the fourth sentence is widely distributed across the scalp, with maxima over left fronto-central and right temporal electrodes. At the third sentence the difference shows maxima over left fronto-central and right tempro-parietal electrodes. These positive differences are the result of a larger decrease in power relative to baseline for the INCOH compared to the COH condition at sentences 3 and 4. In the introduction we linked oscillatory power in the beta frequency range to sentence-level syntactic processing, but pointed out that beta has also been modulated by manipulations of semantic processing. The results reported here might therefore be taken as evidence that discourse-level semantic information can influence either local sentence-level syntactic or semantic processing (or both). We have already argued that our N400 ERP effects are related to sentence-level semantic processing, which is influenced by the discourse-level semantic manipulation. A correlation analysis revealed no relationship between our beta power effects and the N400 ERP findings, suggesting that differences in beta power between COH and INCOH conditions is more likely related to an influence of discourse-level semantic coherence on local syntactic processing.

There are however alternative explanations for the difference in beta power that should also be considered. Weiss and Mueller (2012) suggest that besides binding (unification) during sentence processing, beta activity might play a role in the processing of action semantics, in memory-related processing, or in attention and the violation of expectations. Indeed motor-related beta activity has been shown when comprehending action-related language (e.g., Moreno et al., 2015; van Elk, van Schie, Zwaan, & Bekkering, 2010). We find it unlikely however that this could explain our beta findings, as we did not explicitly manipulate action-semantic content between our COH and INCOH conditions.

The argument for a relationship between beta oscillations and memory processes is based largely on biophysically realistic computational modelling, showing that increases in beta activity in local cortical circuits have exactly the characteristics that would be necessary to hold information online for extended periods of time (Kopell, Whittington, & Kramer, 2011). In contrast, a beta decrease in left frontal cortex has been linked to improved subsequent memory (Hanslmayr, Staudigl, & Fellner, 2012). It is possible that our COH condition placed higher demands on short-term memory than our INCOH condition because related sentences mean a situation model has to be maintained, whereas in the INCOH condition participants don't necessarily construct a situation model. On the other hand, it is also possible that our INCOH condition placed higher demands on short-term memory than our COH condition because unrelated sentences mean that if participants do attempt to construct a situation model after sentence 2, they may have to maintain more information in order to attempt to combine the sentences than in the COH condition (where this is immediately possible). Each of these possibilities fits with one of the above ideas about the relationship between beta and memory processing. Other findings relating beta to short-term memory during language processing have all argued for a relationship between increased beta power and higher memory demands (Haarmann, Cameron, & Ruchkin, 2002; Meyer et al., 2013; Weiss et al., 2005), but these findings can be explained just as well by a link between beta activity and syntactic unification demands (see Lewis, Wang, & Bastiaansen, 2015 for discussion). In fact, Meyer et al. (2013) explicitly argue that for their long- and short-distance subject-verb agreement dependencies *alpha* activity is an index of retention of information in short-term memory, while beta is related to syntactic integration, as it only appears at the end of the retention interval. Given that there is not yet consensus about the exact relationship between beta oscillations and short-term memory during language comprehension, and that all previous findings from language comprehension explicitly relating beta to short-term memory can be adequately explained in terms of differentially demanding syntactic processing, we think that our results are more straightforwardly explained by linking beta to local sentence-level syntactic processing.

Based on our findings we cannot draw any strong conclusions about the exact mechanism by which discourse-level semantic information influences sentence-level syntactic processing. One very tentative proposal is that at points in a sentence when the language comprehension system becomes aware that the sentence-level meaning being constructed does not fit the wider discourse context, the system attempts a syntactic reanalysis in order to try to make the new information fit. This syntactic reanalysis would result in a disruption of local syntactic processing, leading to a decrease in beta power. If this occurred at multiple words within each sentence that would explain why the effect is present in our word averaged data. Our initial hypothesis was that gamma power would be modulated when sentence-level semantic processing was influenced by discourse-level semantic coherence, but we do not observe any gamma effects in the present study.

As we have argued above, beta has also been linked to manipulations of sentence-level semantic processing, and this explanation of the findings cannot be fully ruled out.

The suggested link between beta activity on the one hand, and attention and expectancy violation on the other, is based on a broader proposal linking beta to the maintenance or change of the current cognitive processing set (Engel & Fries, 2010). These ideas have recently been made more explicit for the case of sentence-level language comprehension (Lewis et al., 2015), where beta increases have been linked to the active maintenance of the brain network configuration responsible for the representation and construction of the current sentence-level meaning, while beta decreases have been linked to a change in the underlying network configuration when the system prepares for a new mode of processing. It is possible that for the INCOH condition at sentences 2 to 4 there are points in the sentence at which it becomes clear to the language comprehension system that the input does not fit coherently into the wider discourse or situation model. In such cases, the system might use these as cues, indicating that the current mode of processing (and thus the current sentence-level meaning) must change, and this would result in the observed lower beta power in the INCOH compared to the COH condition. It should be noted that this account of the findings in terms of maintenance/change of processing set is not incompatible with the account in terms of disrupted local syntactic or semantic processing, because as we have argued elsewhere (Lewis et al., 2015), if local syntactic processing is disrupted this should also act as a cue to the system that the current mode of processing has to change. In this experiment we have not explicitly tried to disentangle these two accounts, and thus while we note that both accounts can adequately explain the beta findings, we prefer to interpret them in relation to the more specific link between beta and local sentence-level syntactic processing that is influenced by manipulating discourse-level semantic coherence.

One way we may have been able to shed more light on this issue is by looking into the P600 ERP component as a marker of syntactic integration (although it has also been shown that the P600 is not exclusively related to syntactic processing; e.g., van de Meerendonk, Kolk, Vissers, & Chwilla, 2010), but a limitation of our approach to the analysis (averaging over words within a sentence) is that only time points between 0 and 500 ms (the SOA) relative to word onset contain non-redundant information. This means that it would not have made sense to investigate the typical P600 time window.

A question that arises concerns why this beta effect is not (descriptively) found before sentence 3 and is only statistically significant at sentence 4, while the behavioural results clearly show that in the INCOH condition the stories already become incoherent at the second sentence. One possibility is that participants have to read to the end of the second sentence before they become aware of the breakdown in coherence in the INCOH condition (in the rating task responses are provided after each sentence). On the other hand, we do observe an N400 effect at the second sentence (Figure 2.1), indicating that participants' brains are already sensitive to our discourse-level semantic manipulation during the second sentence. A more likely explanation is that the lack of coherence between sentences in the INCOH condition simply becomes clearer as one moves from reading sentence 2 to sentence 4. At the second sentence participants presumably notice the lack of coherence, but perhaps continue to try to combine the sentences anyway, whereas by the fourth sentence the lack of coherence is clear because sentences 2, 3 and 4 were all incoherent with the first sentence and with each other. Although we did not test this explicitly, and the interaction between sentence position and condition was not statistically significant, visual inspection of the differences in beta power (Figure 2.2) along with the topographies associated with the N400 effects (Figure 2.1) lend support to this idea, as in both cases the difference seems to be larger and more widespread the later a sentence appears in the stories. As we have outlined earlier, our statistical analyses are likely relatively insensitive to interaction effects due to the poor SNR and due to averaging over all words within each sentence, many of which are likely to not exhibit effects related to our experimental manipulation.

Finally, our beta effect shows a difference over left fronto-central electrodes, and over right temporal and parietal electrodes. Although the spatial resolution with EEG is relatively poor, we may speculate that the left frontal difference is related to differential activation of left inferior frontal cortex, that may be disengaged in the INCOH condition when local sentence-level syntactic processing is disrupted (e.g., Hagoort, 2005, 2013; Meyer, Obleser, Kiebel, & Friederici, 2012; Tyler et al., 2011). The right hemisphere differences may be directly related to discourse-level semantic information, as it has been argued (Jung-Beeman, 2005) that the right hemisphere is involved in the recognition of more distant relations between, for example, discourse entities, and also to the activation of broader meaning (perhaps across sentences). In the INCOH condition this network of areas may be disengaged (hence the decrease in beta power) because no situation model can be constructed relating the sentences comprising the stories.

2.4.3 Limitations

This study has two major limitations that should be addressed in future research. First, we were forced to average over all words within each sentence for the two conditions in order to achieve a high enough SNR to observe effects in the data. While we have argued that this approach does retain temporal information in the form of the word averages for each of the four sentences comprising our stories, it does not take full advantage of the EEG data related to the unfolding of each sentence of the stories. Indeed, our original intention was to investigate the evolution of power over the course of the entire discourse, and how that was modulated by discourse-level semantic coherence (Lewis et al., 2012). In addition, we rejected an unusually large number of participants (five) due to poor data quality. We have argued that the relatively poor SNR for our data is due to the unusually long trials, lasting for more than 20 seconds each. This makes it far less likely that participants are able to stay still during each trial, and when they strain to try to do so, likely results in an abundance of electromyographic (EMG) artifacts in the data. It was not possible to remove these using ICA, as they were present for many components, each time with different topographical configurations, often in combination with other non-artifactual activity. Similar future studies should attempt to reduce EMG artifacts by for instance presenting fewer trials for each recording block, and lengthening the inter-sentence interval in combination with explicitly encouraging participants to use the inter-sentence intervals to relax for a moment before keeping still again during the following sentence presentation. The poor SNR for our data may offer an explanation for the absence of any gamma effects related to our discourse-level semantic manipulation. EMG artifacts are especially likely to affect the ability to measure high frequency oscillatory activity (e.g., Hipp & Siegel, 2013), and future studies investigating gamma should also consider analysing oscillatory power at the level of cortical sources, where spatial filters applied during source reconstruction (e.g., Gross et al., 2001) are likely to minimize the effects of EMG activity.

A second limitation is that we do not observe an interaction between sentence position and condition for our beta effect. This means that strictly speaking we cannot make inferential claims about whether the difference between COH and INCOH conditions is present for some sentences (sentences 3 and 4) but not for others (sentences 1 and 2). One potential reason that the interaction is not significant (we hypothesized no difference between COH and INCOH conditions at sentence 1, and a difference at sentences 2 to 4) is that we selected electrodes to test for the interaction based on those electrodes contributing to the significant difference between COH and INCOH conditions at sentence 4 from the output of the cluster-based permutation statistics. Figure 2.2B clearly indicates that the difference between conditions shifts around in space from sentence 3 to sentence 4, and this introduces additional variability into the data. In our opinion, it would not be appropriate to select different electrodes for each sentence to test for an interaction between sentence position and condition, even though that might be most likely to produce the desired result. In addition, there is already a small difference between COH and INCOH conditions at the first sentence (Figure 2.2B), which while it cannot be an effect related to our experimental manipulation (there was no difference between COH and INCOH conditions at the first sentence), does go in the same direction as the effects at sentences 2 to 4, and so probably also contributes to the absence of an interaction between sentence position and condition.

2.4.4 Conclusions

This study shows that discourse-level semantic coherence has an effect on ERP and oscillatory responses related to local sentence-level language comprehension. The semantic processing difficulty that results when sentences are incoherent, leads to an enhanced N400 ERP component. In addition, oscillatory power in the beta frequency range is higher for coherent stories, potentially indicating that discourse-level semantic coherence also affects local syntactic processing. More generally, our beta findings may be related to proposals linking lower beta power to a change in the current cognitive set, where incoherent sentences might result in local sentence-level disruptions in syntactic and/or semantic processing, which in turn provide the language comprehension system with cues indicating that the current mode of processing needs to change. Finally, more attention appears to have been allocated to individual words in each sentence for coherent stories, as indicated by an enhanced visual N1 ERP component. We conclude that discourse-level semantic information is used during local sentence-level language comprehension, and has an effect on electrophysiological signatures of brain activity related to such processing.

Notes

1. Ratings task results presented in Table 2.2 are those for the items included in the final stimulus set. The original rating task (see Hoffman, 2011) included a larger set of stimuli and items not

exhibiting the coherence breakdown at sentences 2, 3 and 4, as indicated by the participant ratings, were excluded from the final set of stimuli used in our experiment.

Chapter 3

Gender Agreement Violations Modulate Beta Oscillatory Dynamics During Sentence Comprehension: A Comparison of Second Language Learners and Native Speakers

Based on:
Lewis, A.G., Lemhöfer, K., Schoffelen, J., & Schriefers, H. (2016). Gender agreement violations modulate beta oscillatory dynamics during sentence comprehension: A comparison of second language learners and native speakers, *Neuropsychologia, 89*, 254-272.

Abstract

For native speakers, many studies suggest a link between oscillatory neural activity in the beta frequency range and syntactic processing. For late second language (L2) learners on the other hand, the extent to which the neural architecture supporting syntactic processing is similar to or different from that of native speakers is still unclear. In a series of four experiments, we used electroencephalography to investigate the link between beta oscillatory activity and the processing of grammatical gender agreement in Dutch determiner-noun pairs, for Dutch native speakers, and for German L2 learners of Dutch. In *Experiment 1* we show that for native speakers, grammatical gender agreement violations are yet another among many syntactic factors that modulate beta oscillatory activity during sentence comprehension. Beta power is higher for grammatically acceptable target words than for those that mismatch in grammatical gender with their preceding determiner. In *Experiment 2* we observed no such beta modulations for L2 learners, irrespective of whether trials were sorted according to objective or subjective syntactic correctness. *Experiment 3* ruled out that the absence of a beta effect for the L2 learners in *Experiment 2* was due to repetition of the target nouns in objectively correct and incorrect determiner-noun pairs. Finally, *Experiment 4* showed that when L2 learners are required to explicitly focus on grammatical information, they show modulations of beta oscillatory activity, comparable to those of native speakers, but only when trials are sorted according to participants' idiosyncratic lexical representations of the grammatical gender of target nouns. Together, these findings suggest that beta power in L2 learners is sensitive to violations of grammatical gender agreement, but only when the importance of grammatical information is highlighted, and only when participants' subjective lexical representations are taken into account.

Keywords: EEG; beta oscillations; grammatical gender; cross-language effects; idiosyncratic lexical representations

3.1 Introduction

The ability to speak a second language has become a valuable and often essential part of everyday life. For adult learners of a second language (L2) some aspects of syntactic processing are extremely difficult, and it is not certain that native-like syntactic processing is attainable (Clahsen & Felser, 2006; Dowens, Guo, Guo, Barber, & Carreiras, 2011; McDonald, 2000; Morgan-Short, Steinhauer, Sanz, & Ullman, 2012). One such aspect is the processing of grammatical gender, which remains problematic at all levels of proficiency (Dewaele & Véronique, 2001; Holmes & Dejean de la Bâtie, 1999; Rogers, 1987). Setting aside arguments about the precise nature of the cognitive system implementing syntactic processing (e.g., Chomsky, 1995; Goldberg, 2003; Jackendoff, 2007), there is general consensus that at a minimum the system has to retrieve lexical representations from long-term memory, and to combine these basic units to form more complex phrase- or sentence-level representations. The neural architecture supporting syntactic processing for native speakers is already well documented (e.g., Friederici, 2002; Hagoort, 2005, 2013; Hickok & Poeppel, 2007). It is not clear however how similar L2 syntactic processing is to syntactic processing in one's native language, and whether or not these involve comparable neural implementations (see e.g., Kotz, 2009; Steinhauer, White, & Drury, 2009).

Measurement methods with high temporal precision are well suited to the investigation of real-time online syntactic processing, and two such methods are Electroencephalography (EEG) and Magnetoencephalography (MEG). Event-related potential/field (ERP/ERF) analyses have proven extremely useful for investigating the timing of various types of linguistic processing (e.g., DeLong, Urbach, & Kutas, 2005; Friederici, 2002; Hagoort & van Berkum, 2007). For native speakers, three main ERP components have been associated with syntactic processing. An early left anterior negativity is sensitive to word category errors and phrase-structure violations (e.g., Friederici, Pfeifer, & Hahne, 1993; Hahne & Friederici, 1999). The left anterior negativity is observed for morphological agreement violations and for various other syntactic violations (e.g., Coulson, King, & Kutas, 1998; Münte, Heinze, & Mangun, 1993). Finally, a P600 is elicited by various syntactic violations (e.g., Osterhout, 1995; Osterhout & Holcomb, 1992, 1993), but also by syntactically complex or ambiguous sentence structures (Kaan, Harris, Gibson, & Holcomb, 2000; Osterhout, Holcomb, & Swinney, 1994).

However, ERP/ERF analyses highlight only part of the rich spatio-spectro-temporal dynamics contained in the EEG/MEG signal (Makeig, Debener, Onton, & Delorme, 2004).

ERPs/ERFs capture time- and phase-locked neural activity by averaging over trials and participants. The standard assumption with ERP analyses is that activity that is not phase-locked to an event should be treated as noise in the recording, but this is not always a valid assumption. Measuring induced oscillatory activity (not phase-locked; e.g., Tallon-Baudry & Bertrand, 1999) in the EEG/MEG can provide additional or complementary information about the underlying cognitive processing. The coupling and uncoupling of functional brain networks is subserved by neural synchronization (e.g., Pfurtscheller & Lopes da Silva, 1999; Singer, 1993; Varela, Lachaux, Rodriguez, & Martinerie, 2001). Synchronous repetitive firing of populations of neurons results in an increased probability that interacting neurons entrain one another in a rhythmic, frequency-specific manner, leading to the creation of functional networks (e.g., König & Schillen, 1991). This supports the integration (or binding) of information distributed over distant neural populations (Gray, König, Engel, & Singer, 1989), and at the same time the segregation of different types of information (represented in networks firing synchronously at different frequencies). We can gain a window onto the dynamic coupling and uncoupling of such functional networks, and how this changes depending on the cognitive task, by measuring frequency-specific oscillatory neural dynamics.

ERP studies have been used to investigate syntactic processing in L2 learners, with mixed results (see Kotz, 2009; Steinhauer et al., 2009). A number of studies have compared ERP findings for the processing of grammatical gender between native speakers and L2 learners (e.g., Gillon Dowens et al., 2011; Gillon Dowens, Vergara, Barber, & Carreiras, 2010; Foucart & Frenck-Mestre, 2012), with the overall conclusion that gender agreement violations are processed similarly by native speakers and L2 learners. On the other hand, L2 learners in these studies were all of relatively high proficiency in their L2, and less proficient L2 learners often have more difficulty with the processing of grammatical gender in their L2 (e.g., Dewaele & Véronique, 2001; Lemhöfer, Schriefers, & Hanique, 2010; Lemhöfer, Spalek, & Schriefers, 2008; Orgassa & Weerman, 2008). A recent study investigated ERP responses to grammatical gender agreement violations in Dutch, comparing Dutch native speakers with German late L2 learners of Dutch who were of approximately intermediate proficiency in their L2 (Lemhöfer, Schriefers, & Indefrey, 2014). They reported a P600 effect for gender agreement violations in the native speaker group but not in the L2 learners of Dutch. More interestingly, when trials were re-sorted according to the L2 learners' subjective representations of correct and incorrect gender agreement, a P600 effect similar to the native speaker group was found. These ERP studies together suggest that while processing of grammatical gender in L2 learners might be dependent on their level of proficiency in the second language, such proficiency effects on syntactic processing may be overestimated by focusing solely on objectively correct and incorrect gender representations. Instead L2 learners' syntactic processing might be comparable to that of native speakers, but carried out based on their (often incorrect) subjective gender representations.

At the same time, there appears to be a link between manipulations of syntactic processing and oscillatory activity in the beta frequency range (13-30 Hz; see Lewis, Wang, & Bastiaansen, 2015 for review). For instance, beta power was higher for centre-embedded relative clauses compared to syntactically less complex right-branching relative clauses (Bastiaansen & Hagoort, 2006), while beta coherence between anterior and posterior electrodes was higher for object-relative clauses compared to syntactically less complex subject-relative clauses (Weiss et al., 2005). Meyer, Obleser, and Friederici (2013) compared long- and short-distance subject-verb agreement dependencies at the point in a sentence where the dependency could be resolved, and found higher beta power in the case of long-distance dependencies. They argued that the higher beta power for long-distance dependencies is related to more demanding syntactic integration. Finally, numerous studies have reported higher beta power at a syntactically correct target word in a sentence compared to syntactically incorrect target words (Bastiaansen, Magyari, & Hagoort, 2010; Davidson & Indefrey, 2007; Kielar, Meltzer, Moreno, Alain, & Bialystok, 2014; Kielar, Panamsky, Links, & Meltzer, 2015). Together these studies suggest that beta power is higher when syntactic processing becomes more challenging, and lower when syntactic processing is disrupted.

Results from syntactic violation studies investigating oscillatory responses to grammatical agreement violations are however mixed. Pérez, Molinaro, Mancini, Barraza, and Carreiras (2012) report higher beta power at a target word for syntactically acceptable sentences compared to sentences containing a grammatical person mismatch between the grammatical subject of the sentence and the target verb. Davidson and Indefrey (2007) presented participants with grammatical number mismatches between subject and target verb in addition to phrase structure violations, but did not find any beta effects for the number agreement violations.

Only one study has investigated oscillatory responses related to syntactic violations in L2 learners (Kielar et al., 2014). They compared sentences with verb tense agreement violations at a target

word to syntactically legal sentences (in addition to sentences containing semantic anomalies). When participants were required to perform an acceptability judgment task, beta power was higher at the target word for syntactically legal sentences. This was the case for both native speakers and L2 learners. Instead, when participants were required to perform a grammaticality judgment task, beta power was again higher at a target word for syntactically legal sentences, but only for the native speaker group. These studies suggest that oscillatory power in the beta frequency range is sensitive to some (person, tense), but possibly not all (number) varieties of grammatical agreement violation. Furthermore, L2 learners also exhibit effects of grammatical agreement violations on beta power, but these effects appear to depend on the task participants are required to perform.
One issue that is not yet clear is exactly which types of syntactic manipulations affect oscillatory activity in the beta frequency range. Since beta power does not appear to be sensitive to violations of grammatical number agreement (Davidson & Indefrey, 2007), it will be important to evaluate the link between oscillatory activity in the beta frequency range and various aspects of syntactic processing. Another outstanding question is whether the relationship between beta activity and syntactic processing is also present for L2 learners, and which factors (e.g., proficiency or task) can influence this link. In this regard, it is important to pay close attention to the role of subjectively compared to objectively correct and incorrect lexical representations (Lemhöfer et al., 2014) at different levels of proficiency, and how this might influence measures of syntactic processing.

In the present study, we address some of these questions by revisiting an existing dataset (*Experiments 1* and *2* below), where participants' EEG was measured while they read syntactically legal sentences and sentences containing a grammatical gender agreement violation, or a grammatical number agreement violation (Lemhöfer et al., 2014). The Gender condition consisted of Dutch singular definite determiner-noun phrases (see Table 3.1 for example stimuli) where the grammatical gender of the determiner and noun either matched (correct trials) or mismatched (incorrect trials). Dutch singular nouns have either neuter or common gender and the corresponding gender marked determiners are het for neuter and de for common gender1. The Number condition consisted of Dutch plural definite determiner-noun phrases, but now the grammatical number of the determiner and noun either matched (correct trials) or mismatched (incorrect trials). Dutch plural nouns should always be preceded by the plural marked determiner de for both neuter and common gender, and thus the neuter singular determiner het together with a plural noun forms a number agreement violation.

Table 3.1 Example materials for *Experiments 1* and *2* and their English translation (in italics).

Condition	*Example Materials*
Gender correct	Hij verzamelde het$_{neu}$ <u>hout</u>$_{neu}$ in een mand en bracht het naar huis.
	He gathered the <u>wood</u> in a basket and brought it home.
Gender incorrect	Ze gebruikte de$_{com}$ <u>hout</u>$_{neu}$ om er een tafel van te maken.
	She used the <u>wood</u> to make a table.
Number correct	Ze zei tegen hem dat de$_{pl}$ <u>hotels</u>$_{pl}$ allemaal al vol zaten.
	She told him that the <u>hotels</u> were all full.
Number incorrect	Het is niet fijn dat het$_{sing}$ <u>hotels</u>$_{pl}$ allemaal duurder zijn geworden.
	It is not nice that the <u>hotels</u> have all become more expensive.

Notes: Assignment of correct and incorrect determiners to sentence frames was counterbalanced across experimental lists. Target nouns are underlined. neu = neuter gender; com = common gender; pl = plural; sing = singular.

A group of Dutch native speakers (*Experiment 1*) read these sentences for comprehension, with occasional comprehension questions after some of the sentences. A time-frequency (TF) analysis of power changes relative to a baseline period immediately prior to the onset of the target word allowed us to test whether oscillatory activity in the beta frequency range is similarly affected for gender agreement processing as it is for other types of grammatical agreement processing. It also allowed us to directly compare how gender and number agreement processing (which was shown to have no effect on beta activity; Davidson & Indefrey, 2007) are related to beta oscillatory activity in the same group of participants. A group of German L2 learners of Dutch (*Experiment 2*) underwent the same procedure, allowing us to test whether effects of gender and number agreement processing on oscillatory activity in the beta frequency range is comparable between native speakers and L2 learners. German L2 learners of Dutch have a tendency to map German neuter gender onto Dutch neuter gender, and to map German feminine and masculine gender onto Dutch common gender (Lemhöfer et al., 2010, 2008). This tendency is particularly strong for cognates between Dutch and German (words that are similar in form and meaning, and have common etymological roots), and can often result in incorrect gender representations for German L2 learners of Dutch (Lemhöfer et al., 2010, 2008). These participants were administered an offline determiner questionnaire after the EEG experiment where they had to provide the correct determiner for each noun they saw in the main experiment. This resulted in a number of trials where subjective and objective correctness mismatched, allowing us to explore how potential beta oscillatory effects linked to syntactic processing are different for participants' subjective and objective gender representations.

3.1.1 Methods common to all experiments

Some details regarding the methods employed are common to *Experiments 1* to *4*, and are therefore described here. EEG data for *Experiments 1* and *2* are from a study by Lemhöfer, Schriefers, and Indefrey (2014). Those authors investigated the effects of gender and number agreement violations on event-related potential (ERP) components. Here we perform a time-frequency (TF) analysis of power with these data in order to investigate induced oscillatory activity related to those experimental manipulations.

3.1.1.1 EEG recordings

Participants were fitted with an elastic cap (Electro-Cap International, Eaton, OH) with electrodes positioned as indicated in Figure 3.1. EEG signals were recorded using 27 passive tin electrodes mounted in the cap and referred to the left mastoid. An additional electrode was placed on participants' right mastoid for re-referencing offline, and a ground electrode was placed on the centre of the forehead. Impedances for these electrodes were kept below 3 kΩ. Additional electrodes were placed on the suborbital and supraorbital ridge of participants' right eye, and on the left and right outer canthi for recording vertical and horizontal EOG activity. EOG electrode impedance was kept below 5 kΩ. EEG and EOG recordings were amplified (8 s time constant; 0.05-30 Hz bandpass filter) and sampled with a frequency of 500 Hz.

3.1.1.2 Data pre-processing

All pre-processing of the EEG data was carried out using Brain Vision Analyzer Version 1.05. For each participant, electrodes were re-referenced to the average of electrodes placed on the left and right mastoid. Data were then segmented from -400 to 1200 ms relative to the onset of the target word. Next, ocular correction was applied to the data from all scalp electrodes using the Gratton & Coles algorithm (Gratton, Coles, & Donchin, 1983) implemented in Brain Vision Analyzer. Data were baseline corrected from -100 to 0 ms relative to target word onset, and a semi-automatic artifact rejection procedure (visual inspection of detected trials; threshold criteria -100 to 100 μV) was used to reject remaining artifactual trials.

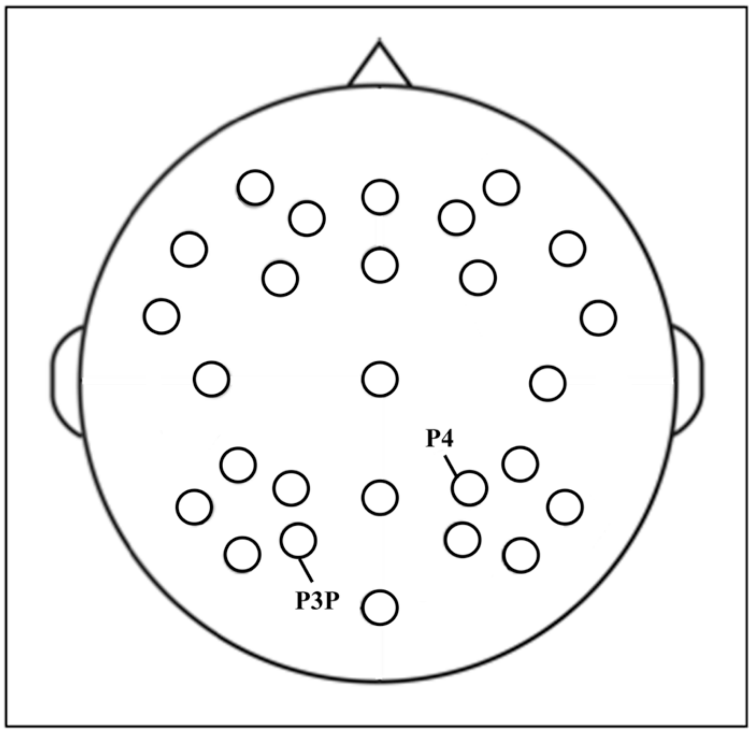

Figure 3.1 Positions of the scalp electrodes in the EEG cap. Representative electrodes used for plotting TF representations of power are indicated.

3.1.1.3 Spectral analysis

All TF, inter-trial coherence (ITC; Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996), and statistical analyses were carried out using the FieldTrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011) running in a Matlab environment (R2012a; Mathworks, Inc.). Fourier spectra of the individual trials were calculated for each participant. In a low-frequency range (2-30 Hz), 400 ms time-smoothing and 2.5 Hz frequency-smoothing windows using a Hanning taper were applied in frequency steps of 1 Hz and time steps of 10 ms.

3.1.1.4 Selection of TF ranges of interest

Single-trial power spectra were computed and averaged for each participant from -200 to 1000 ms relative to the onset of the target word, separately for the Gender and Number conditions in *Experiments 1* and *2* (see *Section 3.2.1* and *3.3.1*), and for the Gender Subjective conditions in *Experiments 3* and *4* (see *Section 3.4.1* and *3.5.1*). This resulted in a TF representation of power for every participant for each condition, irrespective of whether the target word led to a syntactically acceptable agreement relation between the target noun and its preceding determiner. These participant averages were then expressed as a relative change (in dB) from a baseline period between 200 and 0 ms prior to the onset of the target word, separately for each condition. The average TF representation of power over all participants and scalp electrodes was then calculated (separately for each experiment) for visual inspection.

We also computed the ITC for each participant from -200 to 1000 ms relative to the onset of the target word by first normalizing the Fourier spectrum of each trial by its amplitude and then averaging the result across all trials for each participant. The calculation was performed separately for the Gender and Number conditions in *Experiments 1* and *2*, and for the Gender Subjective conditions in *Experiments 3* and *4*, and irrespective of syntactic correctness. This provides a frequency-resolved measure of the degree of trial-to-trial phase consistency over time (Makeig et al., 2004). We used this to distinguish evoked (phase-locked and likely related to the ERP results) from induced (time- but not phase-locked) activity in subsequent TF analyses. Resultant participant-specific ITC values were then averaged over all participants and scalp electrodes (separately for each experiment) for visual inspection.

Next, we selected the same TF ranges of interest for all four experiments based on previous literature and on visual inspection of the TF and ITC data averaged over syntactically correct and incorrect target words (within each condition and experiment separately), over all scalp electrodes, and over participants (Figures 3.2A-F). Our criteria were: 1) a visible increase or decrease in power in the TF representation relative to baseline; 2) only weak (less than 0.15) or no phase-locking visible for the corresponding TF range in the ITC values for any condition in any experiment; 3) good correspondence with previous results in terms of frequency range selected. This resulted in the selection of the following TF ranges of interest to be used in *Experiments 1* to *4* (see black boxes in Figures 3.2A-F): theta: 3-7 Hz and 550-850 ms relative to word onset;

alpha/beta:8-20 Hz and 250-550 ms relative to word onset; beta:12-18 Hz and 650-950 ms relative to word onset.

3.1.1.5 Statistical analyses

Statistical inference was performed using a cluster-based random permutation approach (Maris & Oostenveld, 2007). We used this approach because of its natural handling of the multiple comparisons problem (MCP).

The family-wise error rate is controlled by making use of the spatial autocorrelation in EEG data. In short, a dependent-samples T-test is performed for every data point (mean power in the TF range of interest at each electrode), and these T-values can be interpreted parametrically, to yield P-values that are not corrected for multiple comparisons. A predefined significance level is chosen (in this case, a P-value of 5%, two-tailed) and all data points not exceeding a T-value corresponding to this level are discarded (set to zero). Clusters are calculated from the remaining data points based on their adjacency in space (adjacent electrodes; minimum cluster size of 2 electrodes).

The T-values for all data points in each cluster are then summed to provide cluster-level statistics. A permutation distribution is created by randomly assigning participant averages to one of the two conditions 2000 times, and each time calculating cluster-level statistics as just described. The highest cluster-level statistic from each randomization is entered into the permutation distribution and the cluster-level statistics calculated for the measured data are compared against this distribution. The null hypothesis of exchangeability is rejected at a family-wise error rate corrected confidence level of 5 %, if the largest observed cluster falls in the highest or lowest 2.5th percentile of the randomization distribution.

We compared syntactically correct and incorrect trials separately for the Gender and Number conditions in *Experiment 1*, for the Gender Objective, Gender Subjective, and Number conditions in *Experiment 2*, and for the Gender Subjective conditions in *Experiments 3* and *4*. Mean power values were compared in the selected theta (3-7 Hz; 550-850 ms relative to word onset), alpha/beta (8-20 Hz; 250-550 ms relative to word onset), and beta (12-18 Hz; 650-950 ms relative to word onset) TF ranges, forming clusters only in space.

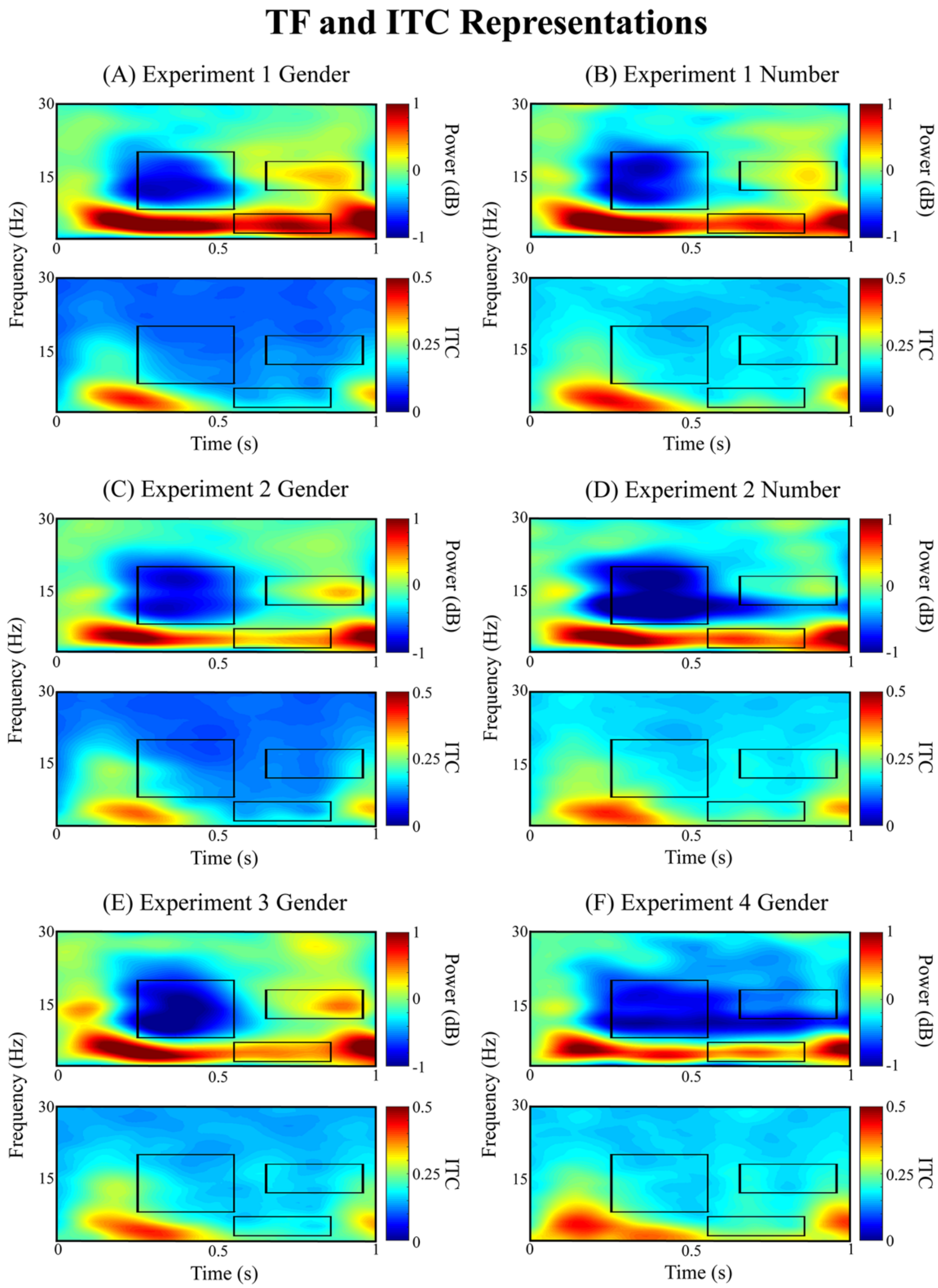

Figure 3.2 TF and ITC representations for all conditions in each experiment. (A) TF (top) and ITC (bottom) representations for the average over all participants, all scalp electrodes, and all target words, irrespective of correctness, for the Gender condition in *Experiment 1*. Black boxes indicate the TF ranges of interest selected for statistical testing. (B) The same as in (A) for the Number condition in *Experiment 1*. (C) The same as in (A) for the Gender (Subjective and Objective are the same in this representation of the data) condition in *Experiment 2*. (D) The same as in (A) for the Number condition in *Experiment 2*. (E) The same as in (A) for the Gender condition in *Experiment 3*. (F) The same as in (A) for the Gender condition in *Experiment 4*.

3.2 Experiment 1

For the first experiment, we hypothesized that if oscillatory activity in the beta frequency range is related to the processing of grammatical gender, then beta power should be higher for syntactically correct target words compared to target words containing a grammatical gender agreement violation. For grammatical number agreement violations, we hypothesized that there should be no difference in beta power between syntactically correct and incorrect target words based on the findings from Davidson and Indefrey (2007).

3.2.1 Methods

3.2.1.1 Participants

Twenty-one native speakers of Dutch took part in the experiment, 19 of whom were included in the final analysis (6 males, 13 females; mean age 23.32 years; SD: 8.08 years). Participants provided informed consent and were paid (10 euros per hour) or equivalently rewarded with course credits for their participation. All participants reported normal or corrected-to-normal vision, and were right handed. None of the participants reported being dyslexic. Two participants were excluded from the final analysis due to recording problems. All Dutch native speakers had experience with other foreign languages (especially English).

3.2.1.2 Stimulus materials

All stimuli consisted of Dutch sentences that were either grammatical or contained a syntactic agreement violation between a target noun and its preceding determiner. Target words were nouns that constituted cognates between Dutch and German.

For target nouns in the Gender condition, 40 Dutch singular nouns were selected from the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995) with compatible grammatical gender between Dutch and German, along with 40 gender-incompatible nouns. Within each of these groups, 20 nouns were of common, and 20 were of neuter gender. Dutch neuter gender nouns were considered compatible with their German counterparts when these were also of neuter gender. Dutch common gender nouns were considered compatible with their German counterparts when these were of either feminine or masculine gender. There was no systematic relationship between the grammatical gender of the target nouns and their natural gender. Two different sentence frames were constructed for every target noun so that participants saw each target noun once with the correct determiner and once with the incorrect determiner (see Table 3.1 for example materials). Within pairs of sentence frames, the sentence structure was as similar as possible up to the target determiner and noun, and these determiner-noun pairs never appeared in sentence-initial or sentence-final position. The occurrence of grammatically correct and incorrect determiner-noun pairs was then counterbalanced across the two sentence frames. A cloze test (see Lemhöfer et al., 2014 for details) revealed that cloze probability for the target noun was below 0.1 for all sentences ($\mu = 0.005$). Inclusion of gender-incompatible cognates at sentence positions other than the target noun was avoided.

For target nouns in the Number condition, 32 Dutch plural nouns were selected from the CELEX database, all with neuter grammatical gender and all gender compatible between Dutch and German. In both Dutch and German there is a single gender-unmarked determiner for plural nouns (de for Dutch and die for German), and hence violations of grammatical number agreement between determiner and noun (the presence of het with a plural target noun) are typically more salient than gender agreement violations for German L2 learners of Dutch. All number agreement violations were constructed by combining a singular determiner (het) with a plural target noun. Two different sentence frames were constructed so that participants saw each target noun once with the correct (de) and once with the incorrect (het) determiner (see Table 3.1 for example materials).

Thirty-two filler sentences were included containing (grammatically correct) plurals of common gender Dutch nouns in target position, all of which were non-cognates between Dutch and German. These sentences were comparable in length and structure to the experimental sentences. Four experimental lists were constructed in order to counterbalance across participants the sentence frame with which a target noun appeared first, and whether a target noun appeared first with a syntactically correct or incorrect determiner. All target nouns appeared once in the first half of the experiment, and once in the second half. Lists were pseudo randomized according to the following restrictions: 1) maximum of three successive correct or incorrect sentences; 2) maximum of three successive sentences containing the same determiner (de or het) directly preceding the target noun; 3) when a sentence was followed by a comprehension question (see *Section 3.2.1.3*), the next sentence was not. Participants were randomly assigned to one of the four experimental lists.

3.2.1.3 Experimental design and procedure

Participants were tested individually in a dimly lit cabin. They were seated in front of a computer monitor and button box with the instruction to read the Dutch sentences presented on the monitor in order to answer occasional questions about them by pressing the right button for a 'yes' response or the left button for a 'no' response. Letters were presented in black on a light grey background using a 24-point sized Arial font type.

Sentences were presented word by word in the centre of the screen. Words were presented for 500 ms, followed by a 300 ms blank screen between words. Each trial began with a fixation cross presented in the centre of the screen for 500 ms, directly followed by a blank screen for 250 ms. The first word of the sentence immediately followed this blank screen. The time between the last word of each sentence and the fixation cross before the next sentence was 1500 ms. Comprehension questions were presented as a whole directly after the last word of a sentence, and remained on the screen until a response was made, or until 10 s had elapsed (this never occurred). Participants were instructed to read all sentences for meaning, and after 10 % of the sentences they answered a yes/no comprehension question. They read a total of 256 sentences (80 correct and 80 incorrect sentences for the Gender condition; 32 correct and 32 incorrect sentences for the Number condition; and 32 fillers), presented in 6 blocks of 44 sentences each. Comprehension questions never followed syntactically incorrect sentences, and half of them required a 'yes' answer. After each block, participants could take a short break, and the first couple of sentences of each new block were additional dummy sentences not included in the analysis. Ten training sentences that were similar to the experimental materials were presented to participants before the experiment. Participants were asked after the EEG experiment whether they noticed anything unusual about the sentences and whether the sentences had been correct, and those who did not notice any grammatical errors were excluded from further analysis. No participants were excluded for this reason in *Experiment 1*. The experimental session lasted between 1.5 and 2 hours in total.

3.2.1.4 Time-frequency analysis

Single-trial power spectra (see *Section 3.1.1.3*) per participant were segmented into syntactically correct and incorrect target word trials for the Gender (Correct: M = 77.63, SD = 2.11; Incorrect: M = 78.68, SD = 1.72) and Number (Correct: M = 31.47, SD = .82; Incorrect: M = 30.89, SD = .97) conditions separately from -200 to 1000 ms relative to word onset. Power spectra were averaged, resulting in participant-specific averages, and these were expressed as a relative change (in dB) from the baseline period between 200 and 0 ms prior to the onset of the target word. Baseline power for each participant was computed separately for the Gender and Number conditions by taking the mean baseline power for syntactically correct and incorrect trials weighted respectively by the number of correct and incorrect trials. This means that when comparing correct and incorrect trials, any power differences observed cannot be due to differences in the level of baseline power (the same power values are used for baseline correction of the correct and incorrect trials). In addition, using a weighted average baseline means that more trials (compared to a condition-specific baseline) contribute to the baseline power estimate, resulting in an improved signal-to-noise ratio. For both the Gender and the Number conditions, this provides us with a measure of the average relative power change from baseline separately for all syntactically correct and incorrect trials.

3.2.2 Results

A separate ERP analysis was performed with the EEG data from *Experiment 1*. Briefly, a P600 effect for both the Gender and Number conditions was observed, with syntactically incorrect target words resulting in a late positive-going deflection in the ERP waveforms compared to syntactically correct target words (see Lemhöfer et al., 2014 for details).

3.2.2.1 Behavioural Results

For the content questions, the mean percentage of errors was 2.32 % for participants included in the final analysis (SD = 3.07 %). Participants were thus paying attention and understood the Dutch sentences they read.

3.2.2.2 TF Results

Figure 3.2A shows the TF representation of power (top) and corresponding ITC values (bottom) for the Gender condition, averaged over all electrodes and trials (regardless of whether they were syntactically correct or incorrect). TF ranges of interest are marked by black boxes. Figure 3.2B shows the same for the Number condition. Statistical comparisons were made between syntactically correct and incorrect trials separately for the Gender and Number conditions for each these TF ranges of interest.

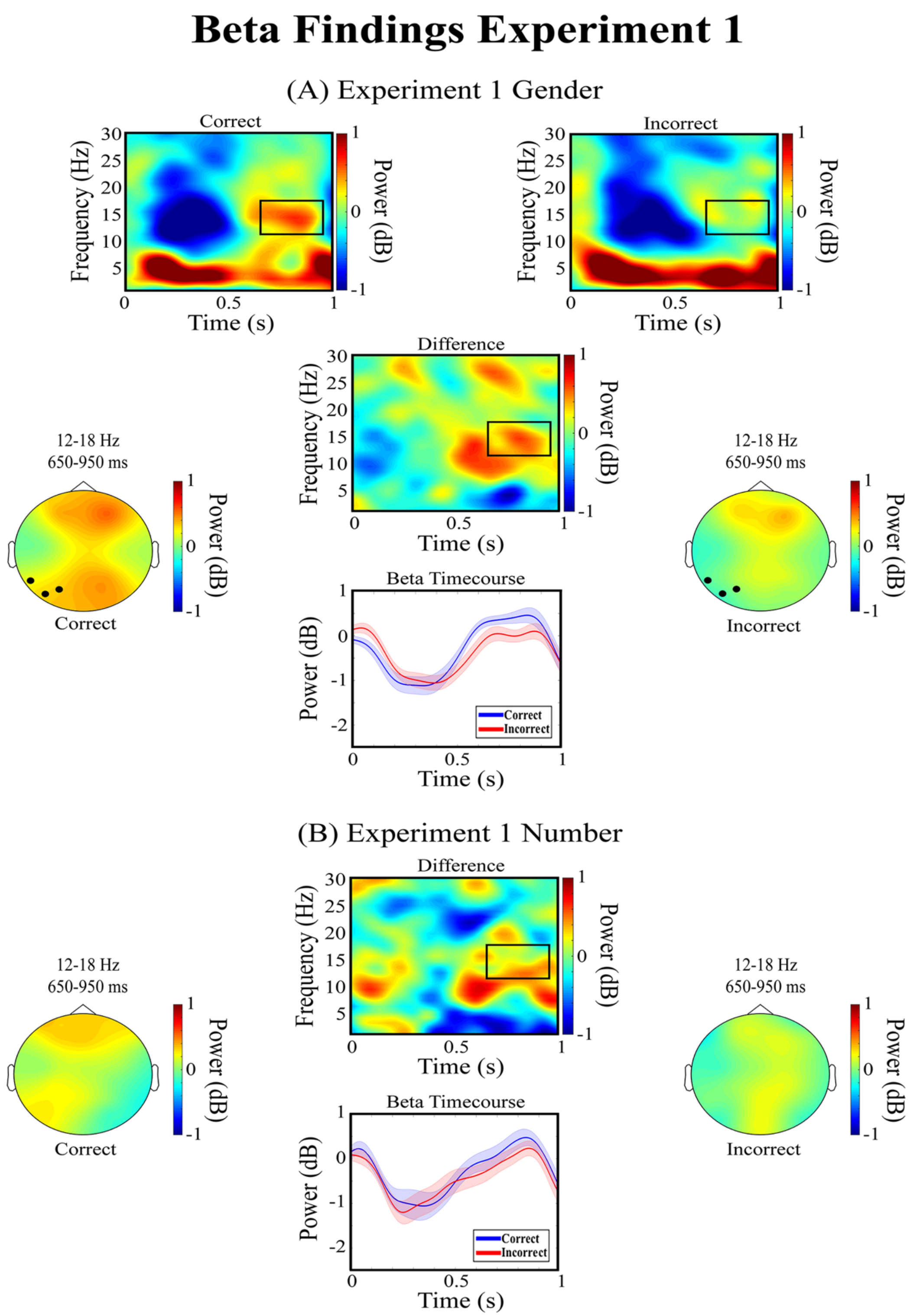

Figure 3.3 Beta TF findings from *Experiment 1*. (A) TF representations of power for correct (top left) and incorrect (top right) target words, and for the difference (correct minus incorrect; middle) at a representative electrode P3P, timecourse of mean beta power (middle bottom) for syntactically correct (blue) and incorrect

(red) target words, as well as scalp distributions of power in the beta TF range of interest (12-18 Hz; 650-950 ms relative to word onset; black boxes in the figure) for correct (bottom left) and incorrect (bottom right) target words in the Gender condition. The difference is statistically significant; black dots indicate electrodes with highest summed T-values during clustering - no inferences can be made based on this information about the spatial extent of the statistically significant effect (Maris & Oostenveld, 2007). (B) TF representation of power (middle top) for the difference between correct and incorrect target words at a representative electrode P3P, along with the timecourse of mean beta power (middle bottom) for syntactically correct (blue) and incorrect (red) target words, and scalp distributions of power in the beta TF range of interest for correct (left) and incorrect (right) target words in the Number condition. The difference is not statistically significant. Shaded regions in the timecourse plots indicate standard error of the mean.

There were no statistically significant differences between syntactically correct and incorrect trials in the alpha/beta TF range of interest for either the Gender or the Number condition. In the beta TF range of interest, there was a significant difference between correct and incorrect trials for the Gender condition ($p = 0.04$). Figure 3.3A shows the TF representation of power for correct (top left) and incorrect (top right) trials, as well as for the difference between the two conditions (middle), at a representative electrode P3P. The figure also shows the timecourse of mean beta power (bottom middle) for syntactically correct (blue) and incorrect (red) trials, as well as the scalp distribution (black dots indicate electrodes contributing to the positive cluster exhibiting highest summed T-values during thresholding – no inferential claims are made regarding the spatial distribution of the significant difference between conditions; see Maris & Oostenveld, 2007) of the mean power in the beta TF range of interest (black box in the TF representations) for correct (bottom left) and incorrect (bottom right) trials. This effect is driven by a late rebound in beta power (after an initial decrease) relative to baseline for correct trials (Figure 3.3A), which is less pronounced for incorrect trials. The scalp distribution of the beta power increase for correct trials shows both frontal and posterior maxima, while for incorrect trials only a weaker frontal maximum is present. There were no statistically significant power differences between syntactically correct and incorrect trials in the beta TF range of interest for the Number condition. For comparison, Figure 3.3B shows the TF representation of power for the difference between correct and incorrect trials (top middle) at a representative electrode P3P, the timecourse of mean beta power (bottom middle) for syntactically correct (blue) and incorrect (red) trials, as well as the scalp distribution of the mean power in the beta TF range of interest for correct (left) and incorrect (right) trials.

In the theta TF range of interest, there was a significant difference between correct and incorrect trials for the Gender condition ($p = 0.032$). Figure 3.5A shows the TF representation of power for the difference between correct and incorrect trials (left) at a representative electrode P4, the timecourse of mean theta power (middle) for syntactically correct (blue) and incorrect (red) trials, as well as the scalp distribution of the mean power in the theta TF range of interest for correct (right top) and incorrect (right bottom) trials (black dots indicate electrodes contributing to the negative cluster exhibiting highest summed T-values during thresholding). This effect is driven by an increase in theta power relative to baseline for both correct and incorrect trials, which is stronger for incorrect trials. The scalp distribution of the theta power increase for correct trials shows a right parieto-occipital maximum, while for incorrect trials it shows a right frontal, in addition to the right parieto-occipital maximum (which is also more pronounced for incorrect trials). There was also a significant difference between correct and incorrect trials ($p = 0.011$) in the theta TF range of interest for the Number condition. Figure 3.4B shows the same as Figure 3.4A, but for the Number condition. The effect is again driven by a larger increase in power for incorrect trials, and the scalp distributions for correct and incorrect trials are very similar to those for the Gender condition, exhibiting a more pronounced right parieto-occipital maximum for incorrect trials and a right frontal maximum for incorrect but not correct trials.

3.2.3 Discussion

The time-frequency analysis of power produced a single statistically significant result in the beta TF range of interest (12-18 Hz; 650-950 ms relative to word onset) for the Gender condition, and a statistically significant result in the theta TF range of interest (3-7 Hz; 550-850 ms relative to word onset) for both the Gender and the Number conditions.

As hypothesized, beta power was higher for syntactically correct than for incorrect target words for the Gender condition, with the largest difference over parietal and occipital electrodes (Figure 3.3A). In the introduction, we linked oscillatory activity in the beta frequency range to syntactic processing, and based on the results reported here, we can add grammatical gender agreement to the list of syntactic features that modulate beta oscillations during sentence comprehension. We also replicated the finding (Davidson & Indefrey, 2007) that grammatical number agreement violations do not result in a modulation of beta power compared to syntactically correct sentences, although in our case, the number mismatch was between a target noun and its preceding determiner, while for Davidson and Indefrey (2007) the mismatch was between a target verb and its preceding grammatical subject.

Theta Findings Experiments 1 and 2

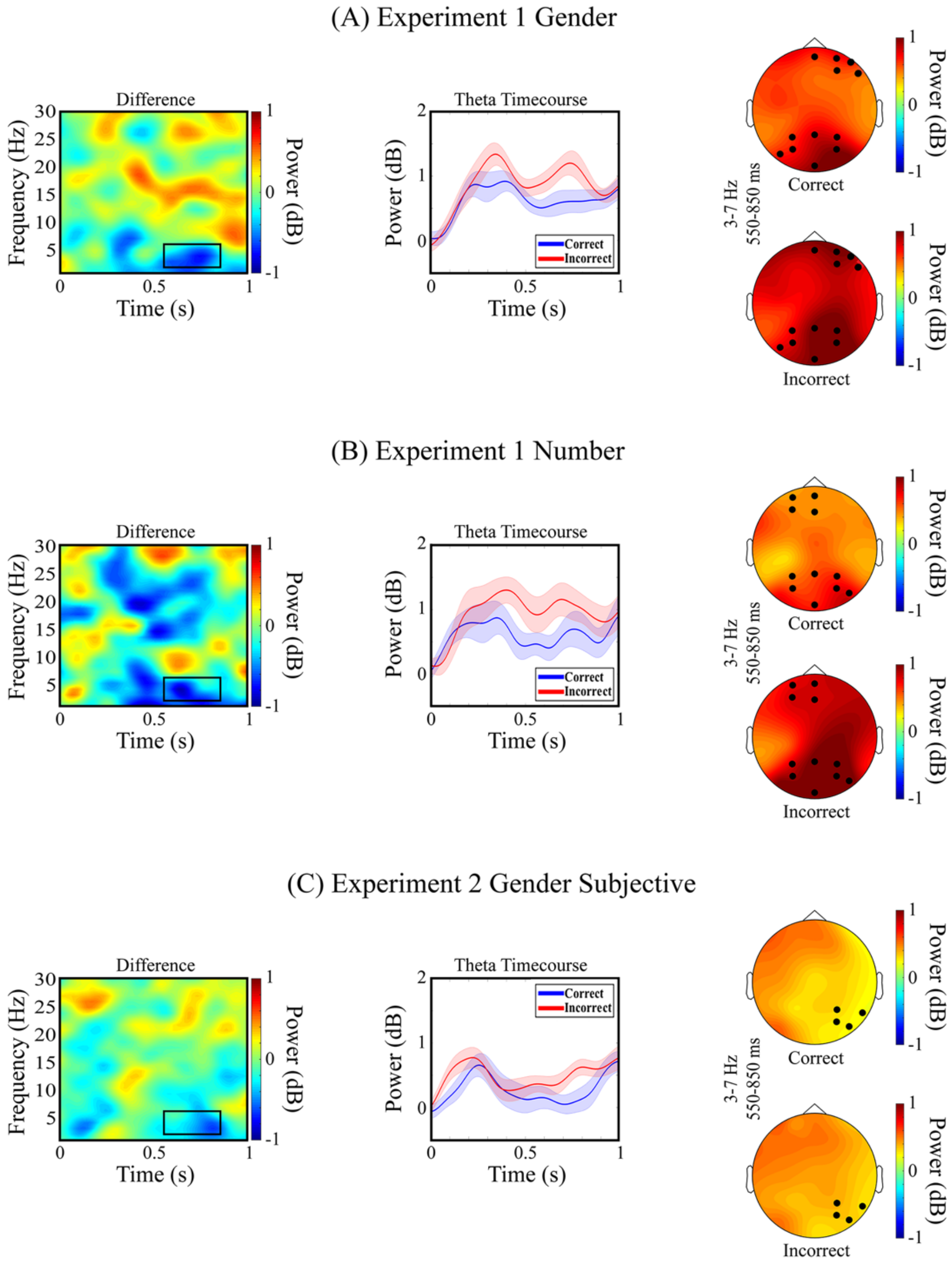

Figure 3.4 Theta TF findings from *Experiments 1* and *2*. (A) TF representation of power (left) for the difference between correct and incorrect target words at a representative electrode P4, along with scalp distributions of power in the theta TF range of interest (3-7 Hz; 550-850 ms relative to word onset; black boxes in the figure) for correct (middle) and incorrect (right) target words in the Gender condition from *Experiment 1*. (B) The same as (A) for the Number condition from *Experiment 1*. (C) The same as (A) for the Gender Subjective condition from *Experiment 2*. All differences are statistically significant. The scale on the power axis for the scalp distributions in each figure matches that for the TF representations of power in the same figure.

For both the Gender and the Number conditions, theta power was higher for syntactically incorrect than for correct target words, with the largest differences over right parieto-occipital and right frontal electrodes (Figures 3.4A and 3.4B). While these were not our main effects of interest, they were not unexpected. Higher theta power has been reported for syntactically incorrect target words involving violations of grammatical number agreement (Bastiaansen, van Berkum, & Hagoort, 2002; Regel, Meyer, & Gunter, 2014), of grammatical gender agreement (Bastiaansen et al., 2002; who did not investigate frequencies higher than alpha), and of grammatical person agreement (Pérez et al., 2012). Based on these findings, it is tempting to relate theta power directly to syntactic processing, but there are a number of syntactic manipulations that do not show theta power modulations (Bastiaansen, Magyari, & Hagoort, 2010; Kielar, Meltzer, Moreno, Alain, & Bialystok, 2014), and moreover, comparable theta effects are most often found for semantic violations during sentence comprehension (Bastiaansen & Hagoort, 2015; Davidson & Indefrey, 2007; Hald, Bastiaansen, & Hagoort, 2006; Kielar et al., 2015). In fact, theta power has explicitly been linked to the retrieval of lexical-semantic information during sentence comprehension (e.g., Bastiaansen, Oostenveld, Jensen, & Hagoort, 2008; Bastiaansen, van der Linden, ter Keurs, Dijkstra, & Hagoort, 2005; Meyer, Grigutsch, Schmuck, Gaston, & Friederici, 2015).

Outside of the domain of language processing, theta power has been implicated in a number of cognitive functions, including working memory (e.g., Gevins, 1997; Kahana, Seelig, & Madsen, 2001; Tesche & Karhu, 2000) and domain general error-detection (e.g., Luu, Tucker, & Makeig, 2004; Luu & Tucker, 2001). Indeed, it has been suggested that theta power might be an index of the building up of a working memory trace of linguistic input during sentence comprehension (Bastiaansen, van Berkum, & Hagoort, 2002). Furthermore, syntactic violations constitute a type of error, and it may not be surprising to find associated oscillatory activity related to domain-general error detection. It is clear that there is a relationship between sentence-level language comprehension and oscillatory power in the theta frequency range, but the precise nature of that relationship (syntactic processing, semantic processing, working memory, etc.) is yet to be determined, and warrants further investigation. An important part of such investigations will be to adequately distinguish induced from evoked theta oscillatory activity, as the latter may simply constitute (at least in part) the frequency domain representation of well-known ERP components (e.g., the N400; Bastiaansen, Mazaheri, & Jensen, 2012), or possibly cortical entrainment to characteristics of the linguistic input (e.g. Luo & Poeppel, 2007; Morillon, Liégeois-Chauvel, Arnal, Bénar, & Giraud, 2012).

3.3 Experiment 2

In *Experiment 2*, German L2 learners of Dutch were tested on the same materials used in *Experiment 1*. The main goal was to test whether the effects of gender agreement processing on oscillatory activity in the beta frequency range found for native speakers is comparable for L2 learners of Dutch. In addition to comparing objectively correct and incorrect determiner-noun gender agreement trials (Gender Objective condition), we re-sorted the trials according to participants' subjective representations of correct and incorrect determiner-noun gender agreement (Gender Subjective condition). This allowed us to explore whether, and how, beta oscillatory effects linked to syntactic processing are different for participants' subjective and objective gender representations.

For the second experiment, we hypothesized that if syntactic processing is similarly affected by gender and number agreement violations in native and L2 speakers, then beta power modulations should be comparable between the two groups. We further hypothesized that if L2 learners rely on their subjective gender representations during online processing, the above similarities should be present only for the Gender Subjective but not for the Gender Objective condition.

3.3.1 Methods

Stimulus materials were the same as those used in *Experiment 1*.

3.3.1.1 Participants

Twenty-nine German learners of Dutch took part in the experiment, 20 of whom were included in the final analysis (4 males, 16 females; mean age 23.1 years; SD: 2.69 years). Participants provided informed consent and were paid (10 euros per hour) or equivalently rewarded with course credits for their participation. All participants reported normal or corrected-to-normal vision, and were right handed. None of the participants reported being dyslexic. Two participants were excluded from the final analysis due to recording problems. A further 7 participants were excluded because of low L2 proficiency and/or lack of critical awareness in Dutch (they did not notice any grammatical errors in the sentences). L2 learners reported speaking other foreign languages besides Dutch. Four participants reported using English more often than Dutch. No other gender marking languages were spoken more often or more proficiently than Dutch. For the full results from a language background questionnaire, see Table 1 in Lemhöfer et al. (2014).

3.3.1.2 Experimental design and procedure

All details were the same as for *Experiment 1* except for the following. After the EEG session, L2 learners were administered an offline questionnaire where all target nouns were listed in random order. They were asked to write down the correct singular definite determiner for each target noun, and to provide a rating for the certainty of their response on a 4-point scale. Their knowledge of plural determiners was also briefly tested by asking them to write down the plural forms of six singular determiner-noun phrases, half of which contained de words and the other half het words. Finally, a language background questionnaire was administered. The experimental session lasted between 2 and 2.5 hours in total.

3.3.1.3 Time-frequency analysis

We analysed the data for the Gender condition in two ways. First, we grouped trials according to objectively correct and incorrect determiner-noun pairs for the target noun as in *Experiment 1* (Gender Objective condition). Second, we re-sorted the trials separately for each participant according to their responses in the offline determiner questionnaire. When participants provided the incorrect determiner for a particular noun, we reversed the objectively correct and incorrect trials so that objectively correct target nouns were labelled as subjectively incorrect, and vice versa (Gender Subjective condition). The resulting grouping provides an indication of participants' brain response to correct and incorrect trials according to their own lexical representations (even if these are objectively inaccurate). For more details of the reasoning behind this approach see Lemhöfer et al. (2014).

Single-trial power spectra (see *Section 3.1.1.3*) per participant were segmented into syntactically correct and incorrect target word trials for the Gender Objective (Correct: M = 77.95, SD = 2.11; Incorrect: M = 78.05, SD = 2.19), Gender Subjective (Correct: M = 77.65, SD = 2.54; Incorrect: M = 78.40, SD = 1.50), and Number (Correct: M = 31.55, SD = .83; Incorrect: M = 31.40, SD = .75) conditions, separately from -200 to 1000 ms relative to target word onset. Power spectra were averaged, resulting in participant-specific averages, and these were expressed as a relative change (in dB) from the baseline period between 200 and 0 ms prior to the onset of the target word. Baseline power for each participant was computed separately for the Gender Objective, Gender Subjective, and Number conditions by taking the mean baseline power for syntactically correct and incorrect trials, weighted respectively by the number of correct and incorrect trials.

3.3.2 Results

A separate ERP analysis was performed with the EEG data from *Experiment 2*. Briefly, a P600 effect, with syntactically incorrect target words resulting in a late positive-going deflection in the ERP waveforms compared to syntactically correct target words, was observed for both the Gender Subjective and Number conditions, but not for the Gender Objective condition (see Lemhöfer et al., 2014 for details).

3.3.2.1 Behavioural Results

For the content questions, the mean percentage of errors was 4.4 % for participants included in the final analysis (SD = 4.66 %). Participants were paying attention and understood the Dutch sentences they read.

The mean error rate for the offline gender questionnaire was 32.62 % (SD = 7.41 %) for participants included in the final analysis. For nouns that are gender incompatible between Dutch and German there were 58.0 % incorrect gender responses, and for gender compatible nouns 7.31 %. No errors were made for the plural definite determiners. As expected, this group of German L2 learners of Dutch have a number of objectively incorrect representations for the grammatical gender of Dutch nouns, and this is especially the case for nouns that are gender incompatible cognates between Dutch and German (e.g., $auto_{com}$, $Auto_{neu}$, car for Dutch, German and English respectively; com denotes common and neu denotes neuter gender marking).

3.3.2.2 TF Results

Figure 3.2C shows the TF representation of power (top) and corresponding ITC values (bottom) for the Gender condition, averaged over all electrodes and trials (regardless of whether they were syntactically correct or incorrect, thus Objective and Subjective are the same in this representation of the data). TF ranges of interest are marked by black boxes. Figure 3.2D shows the corresponding data for the Number condition. Statistical comparisons were made between syntactically correct and incorrect trials separately for the Gender Objective, Gender Subjective, and Number conditions for each these TF ranges of interest.

There were no statistically significant differences between syntactically correct and incorrect trials in the alpha/beta and the beta TF ranges of interest for the Gender Objective, the Gender Subjective, or the Number conditions. We briefly describe the data from the beta TF range of interest for comparison with *Experiment 1*. Figure 3.5A shows the TF representation of power for the difference between correct and incorrect trials (left) in the Gender Objective condition at a representative electrode P3P, the timecourse of mean beta power (middle) for syntactically correct (blue) and incorrect (red) trials, as well as the scalp distribution of the mean power in the beta TF range of interest for correct (right top) and incorrect (right bottom) trials (black dots indicate electrodes contributing to the negative cluster exhibiting highest summed T-values during thresholding). Figures 3.5B and 3.5C show the same for the Gender Subjective and Number conditions respectively. For the Gender Objective and Gender Subjective conditions, it is clear that there is no difference between syntactically correct and incorrect trials. The small positive difference between correct and incorrect trials for the Number condition does not result in any clustering in space.

In the theta TF range of interest, there was a significant difference between correct and incorrect trials for the Gender Subjective condition ($p = 0.03$). Figure 3.4C shows the TF representation of power for the difference between subjectively correct and incorrect trials (left) at a representative electrode P4, the timecourse of mean theta power (middle) for syntactically correct (blue) and incorrect (red) trials, as well as the scalp distribution of the mean power in the theta TF range of interest for correct (right top) and incorrect (right bottom) trials. This effect is driven by an increase in theta power relative to baseline for both correct and incorrect trials, which is more pronounced for incorrect trials. The scalp distribution of the theta power increase for correct trials shows a left frontal maximum, while for incorrect trials this power increase becomes more pronounced and spreads to mid frontal and mid fronto-central electrodes. There were no statistically significant power differences between correct and incorrect trials in the theta TF range of interest for the Gender Objective or the Number conditions.

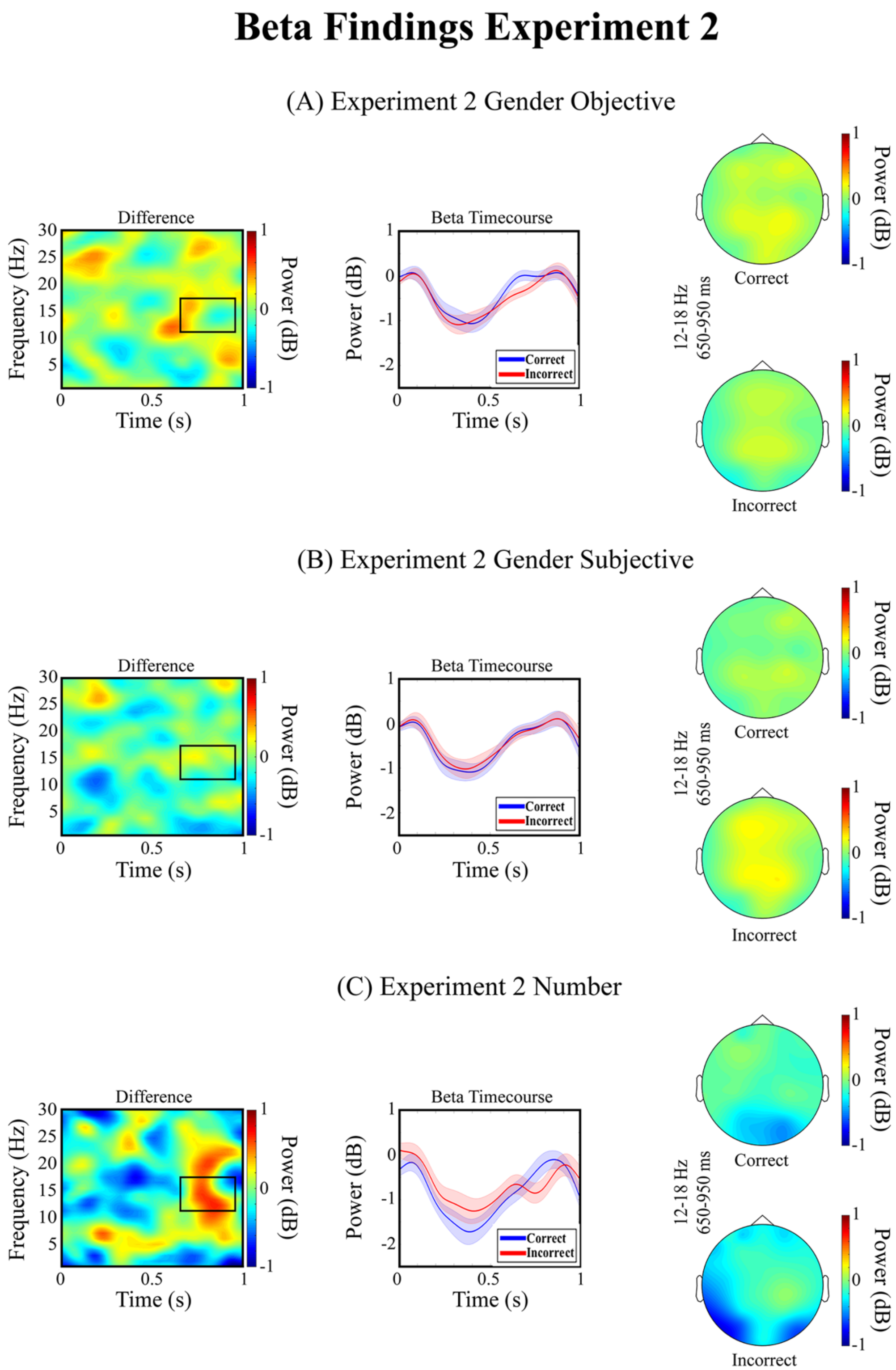

Figure 3.5 Beta TF findings from *Experiment 2*. (A) TF representation of power (left) for the difference between correct and incorrect target words at a representative electrode P3P, timecourse of mean beta power (middle) for syntactically correct (blue) and incorrect (red) target words, as well as scalp distributions of

power in the beta TF range of interest (12-18 Hz; 650-950 ms relative to word onset) for correct (right top) and incorrect (right bottom) target words in the Gender Objective condition. The difference is not statistically significant. (B) The same as (A) for the Gender Subjective condition. The difference is not statistically significant. (C) The same as (A) for the Number condition. The difference is not statistically significant. Shaded regions in the timecourse plots indicate standard error of the mean.

3.3.3 Discussion

The time-frequency analysis of power produced a single statistically significant result in the theta TF range of interest (3-7 Hz; 550-850 ms relative to word onset) for the Gender Subjective condition. There were no differences between correct and incorrect trials in the alpha/beta or the beta TF ranges of interest for any of the conditions in the L2 learners.

For the Gender Subjective condition theta power was higher at syntactically incorrect compared to correct target words, with the largest difference over right parieto-occipital electrodes (Figure 3.4C). The present finding is compatible with our theta results from *Experiment 1*, assuming we accept that participants rely on their subjective lexical representations when processing grammatical gender agreement (Lemhöfer et al., 2014). What these L2 learners of Dutch perceive as incorrectly gender marked target nouns appear to result in similar processing difficulties to those experienced by native speakers, and thus to comparable differences in oscillatory activity in the theta frequency range. We argued in *Experiment 1* that there may be a number of potential roles that theta oscillatory activity plays in sentence-level language comprehension (e.g., syntactic and/or semantic processing, working memory, etc.), and this appears to be the case for L2 learners as for native speakers.

A question that arises is why we did not observe any theta effects for the Number condition, when such effects were clearly observed for the native speakers. Since there were fewer trials in the Number condition than in either of the Gender conditions, it may simply be that there were too few trials in the Number condition to observe any theta effects in the L2 learners. Native speakers likely exhibit more pronounced theta effects than L2 learners, which would explain why theta effects were observed for the Number condition in *Experiment 1* in spite of the low number of trials in that condition.

Like native speakers, L2 learners of Dutch did not show any beta effects for the Number condition (Figure 3.5C). This is exactly as we hypothesized, and we think that in combination with the findings from Davidson and Indefrey (2007), this is a strong case suggesting that the processing of grammatical number agreement does not have an effect on oscillatory activity in the beta frequency range for native speakers, and perhaps also for L2 learners (although we cannot rule out that the absence of beta effects is a result of the relatively low number of trials in the Number condition). It is important to bear in mind that this does not necessarily mean syntactic processing is not affected by number agreement violations. Both native speakers and L2 learners exhibited a P600 effect in the ERP analysis performed on the EEG data from the present experiment (Lemhöfer et al., 2014), indicating that syntactic processing was disrupted by number agreement violations. Instead, it seems that beta oscillations are only sensitive to certain types of syntactic violation, and this does not include violations of grammatical number agreement.

Unlike native speakers however, L2 learners of Dutch did not show beta effects for the Gender Objective condition (Figure 3.5A), and re-sorting trials according to participants' subjective gender representations (Gender Subjective condition; Figure 3.5B) did not produce any statistically significant differences in this frequency range either. One possible explanation for these findings is that L2 learners do not pay close attention to grammatical gender information, adopting a 'shallow' mode of syntactic processing (e.g., Ferreira & Patson, 2007). This is consistent with a number of behavioural studies (e.g., Guillelmon & Grosjean, 2001; Lew-Williams & Fernald, 2010; Scherag, Demuth, Rosler, Neville, & Rocher, 2004), but not with the ERP analysis performed on the EEG data from the present study (Lemhöfer et al., 2014), where both native speakers and L2 learners of Dutch showed a P600 effect. The L2 learners showed a P600 only when trials were re-sorted according to participants' subjective gender representations, and so in our Gender Subjective condition L2 learners appear to be sensitive to grammatical gender information during syntactic processing. Another possible explanation is that oscillatory activity in the beta frequency range is related to syntactic processing in native speakers but not in L2 learners. This possibility strikes us as unlikely, and we will therefore first attempt to rule out two other potential reasons for the absence of a beta effect in our L2 learners.

First, one may suspect that the inclusion of objectively incorrect trials in *Experiment 2* might lead to confusion for our L2 learners, since they saw each target noun twice, once with the objectively correct and once with the objectively incorrect determiner. We selected target nouns for which German L2 learners of Dutch are expected to have difficulty learning the objectively correct grammatical gender. If participants were already uncertain about the correct grammatical gender of the target nouns, seeing these nouns twice during the experiment (with different determiners on each occasion) might have caused them to second-guess whether or not they were correct about the noun's grammatical gender. This could influence participants' neural response, irrespective of whether trials are sorted according to objective or subjective correctness, and could explain the absence of any beta findings for the L2 learners. We investigate this possibility further in *Experiment 3*. Second, it is possible that L2 learners rely less on grammatical gender information during syntactic processing (without ignoring it entirely). Requiring participants to perform a task explicitly focusing their attention on grammatical information might improve the likelihood of observing an effect of gender agreement violations on beta oscillatory activity. We address this possibility in *Experiment 4*.

3.4 Experiment 3

In *Experiment 3*, a new sample of German L2 learners of Dutch were tested on a new set of stimuli, similar to those used in the first two experiments, but this time not containing the Number condition and not containing any objectively incorrect determiner-noun gender agreement trials. In other words, all sentences in this experiment were objectively syntactically correct (note that this is likely more representative of the kind of second language input these participants are accustomed to in their day-to-day lives). Participants completed an offline determiner questionnaire, embedded in other tests of Dutch language skills, approximately 1 week before the main EEG experiment. Trials from the EEG experiment were sorted according to participants' subjective representations of correct and incorrect determiner-noun pairs (Gender Subjective condition), and a time-frequency analysis of power was performed. This allowed us to investigate whether the within-participant repetition of each noun, once with each determiner, in Experiment 2 might have resulted in the absence of any beta findings for the L2 learners. For *Experiment 3*, we hypothesized that if the absence of beta findings in *Experiment 2* was due to the repetition of nouns in correct and incorrect trials, then in the present experiment L2 learners should exhibit beta effects similar to those observed in Experiment 1 for the native speaker group.

3.4.1 Methods

All details regarding the methods employed in this experiment are the same as in *Experiment 1*, apart from those reported below.

3.4.1.1 Participants

Twenty-eight German learners of Dutch took part in the experiment, 20 of whom were included in the final analysis (1 male, 19 females; mean age 23.2 years; SD: 2.17 years). Participants provided informed consent and were paid (10 euros per hour) or equivalently rewarded with course credits for their participation. All participants reported normal or corrected-to-normal vision, and were right handed. EEG data were not recorded for two participants because they made very few errors on the offline determiner questionnaire, indicating that there would likely be very few subjectively incorrect trials for subsequent analyses. Three other participants did not return for the EEG session. These participants were excluded from further analysis. One participant was excluded because they turned out to be dyslexic, and another due to poor performance on the language proficiency questionnaire (see *Section 3.4.1.3* below). One participant was excluded from the final analysis due to recording problems. L2 learners reported speaking other foreign languages besides Dutch. Five participants reported speaking English more proficiently than Dutch and one (different) participant reported speaking English more often than Dutch. No other gender marking languages were spoken more often or more proficiently than Dutch. All results from the language background questionnaire are summarized in Table 3.2.

Table 3.2 Results from Language Background Questionnaire from *Experiment 3*.

	Mean	*SD*	*Range*
Age of first contact with Dutch (years)	20.2	1.5	18-24
Years of experience with Dutch	3.2	1.2	1.5-5
Self-ratings[a]			
How often do you read Dutch literature?	5.4	1.7	1-7
How often do you speak Dutch?	5.8	1.0	4-7
How often do you listen to Dutch radio/watch Dutch TV?	4.0	1.7	1-7
Reading experience in Dutch	5.2	1.2	3-7
Writing experience in Dutch	4.9	1.0	3-7
Speaking experience in Dutch	5.2	1.0	4-7
Mean Dutch experience (mean previous 3)	5.1	0.8	4-7

Notes: [a]Self-ratings were given on a scale from 1 (*low/rarely*) to 7 (*high/very often*).

3.4.1.2 Stimulus materials

All stimuli consisted of grammatically correct Dutch sentences. Target words were nouns that constituted cognates between Dutch and German.

For target nouns, 68 Dutch singular nouns were selected from the CELEX database (Baayen et al., 1995) with compatible grammatical gender marking between Dutch and German. For each gender compatible target noun, a gender incompatible target noun was selected, matched for log frequency (compatible: μ = 1.31; SD = 0.54; incompatible: μ = 1.33; SD = 0.55), orthographic overlap between Dutch and German (compatible: μ = 86.2 %; SD = 14.6 %; incompatible: μ = 87.1 %; SD = 17 %), and word length in letters (compatible: μ = 5.72; SD = 1.66; incompatible: μ = 5.74; SD = 1.58). The target nouns of each pair had the same grammatical gender in Dutch. Within each of the compatible and incompatible target noun groups, 34 were of common gender (requiring the determiner *de*) and 34 were of neuter gender (requiring the determiner *het*). About 60 % of the nouns were the same as those used in *Experiments 1* and *2*. When that was not the case, this was because those words could not be used with the current stimulus requirements.

Two different sentence frames were constructed for every target noun so that participants saw each target noun once with a definite determiner (*het* for neuter gender and *de* for common gender) and once with an indefinite determiner (*een*, not marked for grammatical gender; see Table 3.3 for example materials). An online web based cloze test revealed that mean cloze probability for the target noun was 0.012.

Ten warmup sentences were included, comparable in length and structure to the experimental sentences. Four experimental lists were constructed in order to counterbalance across participants the sentence frame with which a target noun appeared first, and whether a target noun appeared first with a definite or an indefinite determiner. Pairs of matched (compatible-incompatible) target nouns appeared in a sentence frame with the same determiner type in the same block. Lists were pseudo randomized according to the following restrictions: 1) maximum of three successive definite determiner sentences that are either gender compatible or incompatible between Dutch and German; 2) maximum of three successive sentences containing the same determiner (*de*, *het*, or *een*) directly preceding the target noun. Participants were randomly assigned to one of the four experimental lists.

Table 3.3 Example materials for *Experiments 3* and *4* and their English translation (in italics).

Condition	*Example Materials*
Definite determiner	De moeder kon zich het$_{neu}$ lied$_{neu}$ niet meer herinnerin.
	The mother could not remember the song.
Indefinite determiner	De straatmuzikant speelde een lied$_{neu}$ dat ze kende.
	The street musician played a song that he knew.

Notes: Assignment of definite and indefinite determiners to sentence frames was counterbalanced across experimental lists. Target nouns are underlined. Only a neuter gender example is presented here but common gender target nouns were also present in the definite determiner condition. neu = neuter gender.

3.4.1.3 Experimental design and procedure

The experiment was split into two sessions. In order to avoid as much as possible drawing participants' attention to determiners during the EEG session, the behavioural session always took place approximately 1 week (range 5 to 9 days) before the EEG session. In the behavioural session, participants were administered an offline questionnaire where all target nouns were listed in random order. Participants were asked to write down the correct singular definite determiner for each target noun, and to provide a rating for the certainty of their response on a 4-point scale. They also completed a LexTALE vocabulary test (Lemhöfer & Broersma, 2012) in Dutch, and the same language background questionnaire as administered in *Experiment 2*. The behavioural session lasted between 0.5 and 1 hours in total.

In the EEG session, participants read a total of 282 sentences (136 definite and 136 indefinite determiner sentences for the Gender condition; and 10 warmup sentences), presented in 6 blocks of 47 sentences each. Participants were instructed to read all sentences for meaning, and after 20 % of the sentences they answered a yes/no comprehension question. After each block, participants could take a short break, and the first couple of sentences of each new block were always warmup sentences not included in the analysis. The EEG session lasted between 1.5 and 2 hours in total.

3.4.1.4 Time-frequency analysis

For each participant, target nouns were only included in the final analysis if the following two criteria were met: 1) in the offline determiner questionnaire participants provided the objectively correct determiner for a noun that was gender compatible between Dutch and German; and 2) in the same questionnaire participants provided the objectively incorrect determiner for the corresponding matched noun (see *Section 3.4.1.2*) that was gender incompatible between Dutch and German. In this way, we ensured an equal number of subjectively correct and incorrect determiner-noun pairs in the analysis (Gender Subjective condition). For our analyses, subjectively incorrect trials consisted of target nouns meeting the second criterion above, and target nouns meeting the first criterion were analysed as subjectively correct trials. The resulting grouping provides an indication of participants' brain response to correct and incorrect trials according to their own lexical representations (even if these are objectively inaccurate).

Single-trial power spectra per participant were segmented into syntactically correct and incorrect target word trials for the Gender Subjective (Correct: M = 39.9, SD = 9.72; Incorrect: M = 39.5, SD = 9.86) condition from -200 to 1000 ms relative to target word onset. Power spectra were averaged, resulting in participant-specific averages, and these were expressed as a relative change (in dB) from the baseline period between 200 and 0 ms prior to the onset of the target word. Baseline power for each participant was computed by taking the mean baseline power for syntactically correct and incorrect trials, weighted respectively by the number of correct and incorrect trials.

3.4.2 Results

A separate ERP analysis was performed with the EEG data from *Experiment 3*. Briefly, there was no evidence for a significant ERP effect for the Gender Subjective condition when comparing syntactically correct and incorrect target nouns (the full ERP results of *Experiment 3* will appear in a separate publication, together with the ERP results from *Experiment 4*).

3.4.2.1 Behavioural Results

For the content questions, the mean percentage of errors was 6.34 % for participants included in the final analysis (SD = 3.25 %). Participants were paying attention and understood the Dutch sentences they read.

The mean error rate for the offline gender questionnaire was 35.91 % (SD = 7.35 %) for participants included in the final analysis. For nouns that are gender incompatible between Dutch and German there were 64.79 % incorrect gender responses, and for gender compatible nouns 7.07

%. As expected, this group of German L2 learners of Dutch have a number of objectively inaccurate representations for the grammatical gender of Dutch nouns, and this is especially the case for nouns that are gender incompatible cognates between Dutch and German.

3.4.2.2 TF Results

Figure 3.2E shows the TF representation of power (top) and corresponding ITC values (bottom) for the Gender Subjective condition, averaged over all electrodes and trials (regardless of whether they were syntactically correct or incorrect). TF ranges of interest are marked by black boxes. Statistical comparisons were made between syntactically correct and incorrect trials for the Gender Subjective condition.

There were no statistically significant differences between syntactically correct and incorrect trials in the theta, the alpha/beta or the beta TF ranges of interest. We briefly describe the data from the beta TF range of interest for comparison with *Experiments 1* and *2*. Figure 3.6A shows the TF representation of power for the difference between correct and incorrect trials (middle top) in the Gender Subjective condition at a representative electrode P3P, the timecourse of mean beta power (middle bottom) for syntactically correct (blue) and incorrect (red) trials, as well as the scalp distribution of the mean power in the beta TF range of interest for correct (left) and incorrect (right) trials. While it may appear from the scalp distributions that beta power is higher for correct than for incorrect trials, the difference does not exhibit any clustering in space.

3.4.3 Discussion

The time-frequency analysis of power did not produce statistically significant results in any of the TF ranges of interest. In *Experiment 2* we observed effects in the theta TF range of interest for the Gender Subjective condition, but not for the Number condition. We argued that there were too few trials in the Number condition to observe possible effects of syntactic violations with L2 learners. In this experiment we do not find any theta effects for the Gender Subjective condition, and we think the same explanation may hold. Only trials where subjective correctness corresponded to what we expected based on whether or not the target noun exhibited compatible or incompatible gender between Dutch and German were included in the analysis. This resulted in about half the number of trials in the Gender Subjective condition as there were in *Experiment 2*, and a comparable number to those used in the Number condition in that experiment.

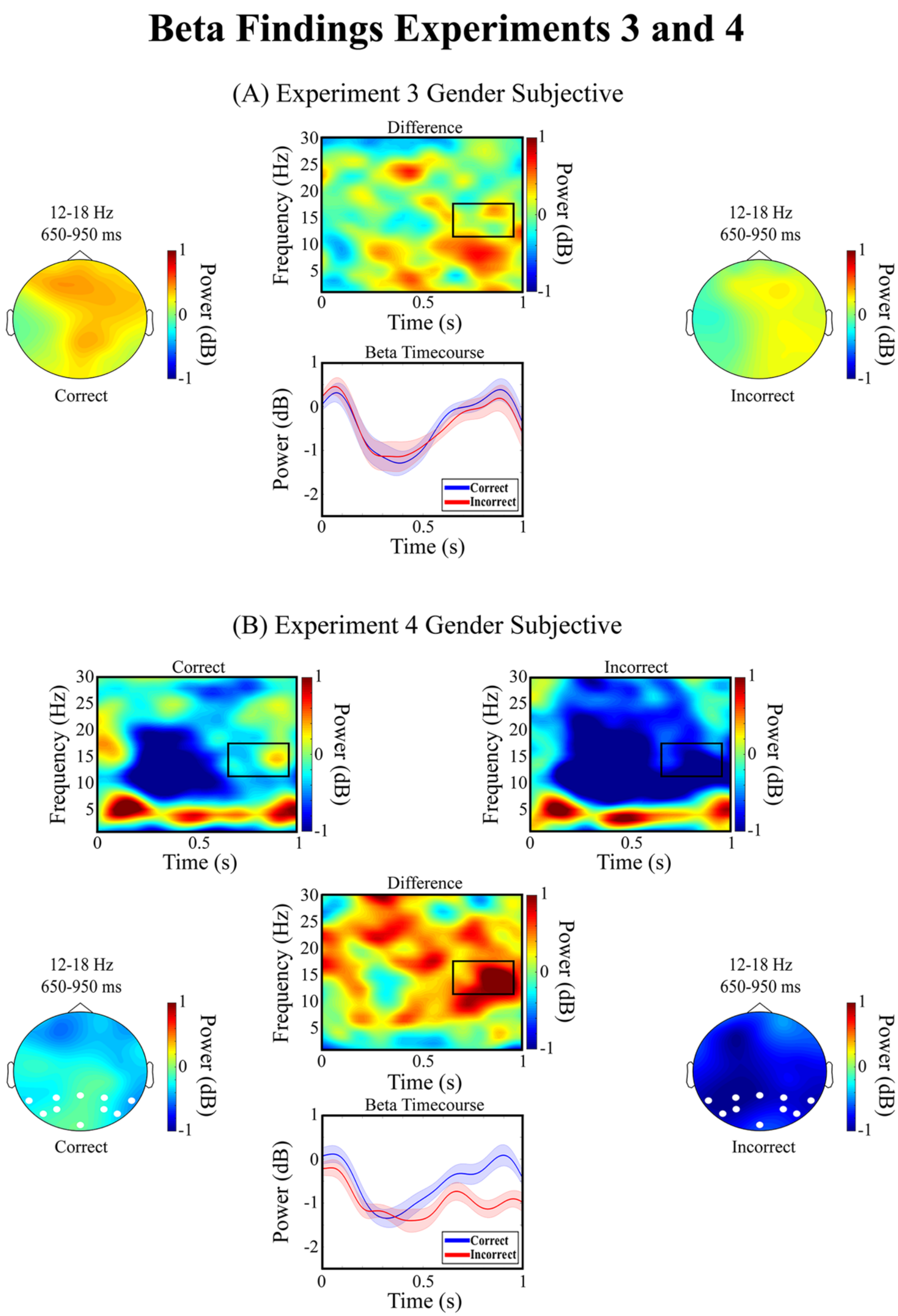

Figure 3.6 Beta TF findings from *Experiments 3* and *4*. (A) TF representation of power (middle top) for the difference between correct and incorrect target words at a representative electrode P3P, timecourse of

mean beta power (middle bottom) for syntactically correct (blue) and incorrect (red) target words, as well as scalp distributions of power in the beta TF range of interest (12-18 Hz; 650-950 ms relative to word onset; black boxes in the figure) for correct (left) and incorrect (right) target words in the Gender Subjective condition from *Experiment 3*. The difference is not statistically significant. (B) TF representations of power for correct (top left) and incorrect (top right) target words, and for the difference (correct minus incorrect; middle) at a representative electrode P3P, timecourse of mean beta power (bottom middle) for syntactically correct (blue) and incorrect (red) target words, as well as scalp distributions of power in the beta TF range of interest for correct (bottom left) and incorrect (bottom right) target words in the Gender Subjective condition from *Experiment 4*. The difference is statistically significant. White dots indicate electrodes with highest summed T-values during clustering. Shaded regions in the timecourse plots indicate standard error of the mean.

All the sentences used in this experiment were objectively syntactically correct. This means we can rule out L2 learners' confusion due to the inclusion of objectively incorrect gender marked nouns as the reason for the absence of beta effects for the Gender Subjective condition in *Experiment 2*. We cannot rule out a lower number of trials as a reason for the absence of any beta effects in the present experiment, but we will return to this point later in the discussion section of *Experiment 4*. Participants' task in the present experiment was to answer yes/no comprehension questions after 20 % of the sentences they read. If L2 learners do make less use of grammatical gender information during syntactic processing, we might be more likely to observe beta effects when participants are required to perform a task explicitly focusing their attention on grammatical information. In *Experiment 4* we repeated *Experiment 3*, but now participants were required to perform a grammaticality judgment task, providing answers about whether or not a sentence was grammatical, after each sentence they read.

3.5 Experiment 4

In *Experiment 4* German L2 learners of Dutch were tested on the same materials used in *Experiment 3*. Participants were now asked to judge the correctness of the determiners in each sentence they read. We hypothesized that using a task explicitly requiring L2 learners to pay attention to grammatical processing should result in beta findings similar to those in Experiment 1.

3.5.1 Methods

All details regarding the methods employed in this experiment are the same as in *Experiment 3*, apart from those reported below.

3.5.1.1 Participants

Twenty-three German learners of Dutch took part in the experiment, 19 of whom were included in the final analysis (3 males, 16 females; mean age 22.74 years; SD: 1.82 years). Participants provided informed consent and were paid (10 euros per hour) or equivalently rewarded with course credits for their participation. All participants reported normal or corrected-to-normal vision, and were right handed. EEG data were not recorded for three participants because they made very few errors on the offline determiner questionnaire (see *Section 3.5.1.2* below), indicating that there would likely be very few subjectively incorrect trials for subsequent analyses. One participant was excluded because it turned out they had not learned German from birth. L2 learners reported speaking other foreign languages besides Dutch. Eight participants reported speaking English more proficiently than Dutch and three of those participants reported speaking English more often than Dutch. No other gender marking languages were spoken more often or more proficiently than Dutch. The results from the language background questionnaire are summarized in Table 3.4.

Table 3.4 Results from Language Background Questionnaire from *Experiment 4*.

	Mean	*SD*	*Range*
Age of first contact with Dutch (years)	20.1	1.4	16-23
Years of experience with Dutch	2.8	1.8	1-6.5
Self-ratings[a]			
How often do you read Dutch literature?	4.8	1.5	1-7
How often do you speak Dutch?	5.0	1.6	2-7
How often do you listen to Dutch radio/watch Dutch TV?	3.1	1.9	1-7
Reading experience in Dutch	4.7	1.0	3-7
Writing experience in Dutch	4.4	1.2	3-7
Speaking experience in Dutch	4.9	1.3	3-7
Mean Dutch experience (mean previous 3)	4.7	1.0	3-7

Notes: [a]Self-ratings were given on a scale from 1 (*low/rarely*) to 7 (*high/very often*).

3.5.1.2 Experimental design and procedure

Testing took place in one experimental session, with all behavioural tests administered after the EEG recording session. Again, participants were instructed to read all sentences for meaning, but additionally, they had to answer a yes/no question presented after each sentence, indicating whether or not the definite determiners in the sentence had all been correct. If they indicated that the sentence was ungrammatical they were asked to report which noun was preceded by an incorrect determiner. In the behavioural session, participants were administered the same offline determiner questionnaire, Dutch LexTALE vocabulary test (Lemhöfer & Broersma, 2012), and language background questionnaire as in *Experiment 3*. The experimental session lasted between 2 and 3 hours in total.

3.5.1.3 Time-frequency analysis

Trials were only included in the analysis if participants' online grammaticality judgments during the EEG session matched their judgments in the offline determiner questionnaire.

Single-trial power spectra per participant were segmented into syntactically correct and incorrect target word trials for the Gender Subjective (Correct: M = 32.5, SD = 8.58; Incorrect: M = 32.6, SD = 8.57) condition from -200 to 1000 ms relative to target word onset. Power spectra were averaged, resulting in participant-specific averages, and these were expressed as a relative change (in dB) from the baseline period between 200 and 0 ms prior to the onset of the target word. Baseline power for each participant was computed by taking the mean baseline power for syntactically correct and incorrect trials, weighted respectively by the number of correct and incorrect trials.

3.5.2 Results

A separate ERP analysis was performed with the EEG data from *Experiment 4*. Briefly, subjectively incorrect target nouns gave rise to a positivity at left ($p = .001$) and right ($p = .006$) posterior electrode sites between 500 and 1000 ms (a P600 effect) in a quadrant-style statistical analysis, when compared to subjectively correct target nouns (the full ERP results of *Experiment 4* will appear in a separate publication, together with the ERP results from *Experiment 3*).

3.5.2.1 Behavioural Results

For the grammaticality judgments, the mean percentage of 'no' responses for target nouns that were gender incompatible between Dutch and German was 54.36 % for participants included in the final analysis, while for gender compatible target nouns it was 6.99 %. As expected, this group of German L2 learners of Dutch have a number of objectively inaccurate representations for the grammatical gender of Dutch nouns, and this is especially the case for nouns that are gender incompatible cognates between Dutch and German.

The mean error rate for the offline gender questionnaire was 35.64 % (SD = 8.69 %) for participants included in the final analysis, while the mean degree of consistency between responses in the offline questionnaire and the online grammaticality judgment task was 80.57 % (SD = 7.67 %). This suggests that participants' incorrect representations are relatively stable, and that judgments are comparable for the two tasks.

3.5.2.2 TF Results

Figure 3.2F shows the TF representation of power (top) and corresponding ITC values (bottom) for the Gender Subjective condition, averaged over all electrodes and trials (regardless of whether they were syntactically correct or incorrect). TF ranges of interest are marked by black boxes. Statistical comparisons were made between syntactically correct and incorrect trials for the Gender Subjective condition.

There were no statistically significant differences between syntactically correct and incorrect trials in the theta or the alpha/beta TF ranges of interest. In the beta TF range of interest, there was a significant difference between correct and incorrect trials for the Gender Subjective condition ($p = 0.001$). Figure 3.6B shows the TF representation of power for correct (top left) and incorrect (top right) trials, as well as for the difference between the two conditions (middle), at a representative electrode P3P. The figure also shows the timecourse of mean beta power (bottom middle) for syntactically correct (blue) and incorrect (red) trials, as well as the scalp distribution of the mean power in the beta TF range of interest for correct (bottom left) and incorrect (bottom right) trials (white dots indicate electrodes contributing to the positive cluster exhibiting highest summed T-values during thresholding). This effect is driven by a prolonged decrease in beta power relative to baseline for incorrect trials, while beta power for correct trials exhibits a late rebound, returning to baseline levels (Figure 3.6B). The scalp distribution of the beta power decrease for incorrect trials shows a maximum over most of the left hemisphere, which is not present for correct trials.

3.5.3 Discussion

The time-frequency analysis of power produced a single result in the beta TF range of interest (12-18 Hz; 650-950 ms relative to word onset) for the Gender Subjective condition. There were no statistically significant differences between correct and incorrect trials in the theta or the alpha/beta TF ranges of interest.

As hypothesized, requiring L2 learners to perform a task that explicitly focused their attention on the grammar of their second language, resulted in beta effects comparable to those found in *Experiment 1* for the native speakers. Beta power was higher for syntactically correct than for incorrect target words for the Gender Subjective condition, with the largest difference over the left hemisphere (Figure 3.6B). It therefore appears that the presence of beta oscillatory effects related to grammatical gender processing in L2 learners is dependent on the extent to which participants are required to explicitly focus on grammatical information.

The number of trials per condition was again relatively low compared to that in the Gender Subjective condition in *Experiment 2*, and was comparable to the number of trials per condition for the Number condition in *Experiment 2* and the Gender Subjective condition in *Experiment 3*. That may again be the reason we did not observe any theta effects in this experiment, despite participants being required to explicitly focus their attention on grammatical information. This could be an indication that the theta effects observed in *Experiments 1* and *2* are not directly related to the processing of syntactic information, but this remains speculative and warrants further research. However, we can rule out the low number of trials as the reason for the absence of beta effects in *Experiment 3*, since we do find beta effects here with a comparable number of trials.

3.6 General Discussion

In the present study we conducted oscillatory analyses on EEG data from a series of four experiments, to investigate how beta power is modulated by violations of syntactic gender agreement in native speakers and L2 learners. Table 3.5 provides an overview of all statistically significant TF and ERP results (ERP results for *Experiments 3* and *4* will be reported in full in a separate publication). One striking aspect of the data is the similarity across experiments of the TF representations of power when averaged over all electrodes, participants, and trials irrespective of correctness of the target word (Figures 3.2A to F). In all conditions, for both native speakers and L2 learners, we observe an early theta power increase relative to baseline, which the ITC representation indicates is likely phase-locked activity, probably related to early ERP components (see e.g., Bastiaansen, Mazaheri, & Jensen, 2012). There is also an early alpha/beta power decrease and a later theta power increase, neither of which are strongly phase-locked to the target word, and so are likely related to induced oscillatory activity. A late increase in beta power is also present, but is very weak in the Number condition in *Experiment 2* (Figure 3.2D) and absent in the Gender Subjective condition in *Experiment 4* (Figure 3.2F). In both of these cases, the earlier alpha/beta power decrease is prolonged in time (especially in the alpha frequency range) compared to all other TF representations.

Table 3.5 Summary of TF and ERP results from all experiments.

Experiment	*Beta*	*Theta*	*P600*
Experiment 1 - Gender	+	-	-
Experiment 1 - Number	=	-	-
Experiment 2 - Gender Objective	=	=	=
Experiment 2 - Gender Subjective	=	-	-
Experiment 2 - Number	=	=	-
Experiment 3 - Gender Subjective	=	=	=
Experiment 4 - Gender Subjective	+	=	-

Notes: Results columns indicate direction of effect for power or amplitude (+ indicates correct > incorrect; - indicates correct < incorrect; = indicates no statistically significant difference).

Explicitly focusing participants' attention on grammatical information by requiring a grammaticality judgment after every sentence appears to have resulted in a prolonged alpha (and to some extent beta) power decrease (Figure 3.2F; although this was not formally tested and so is only descriptive). This is consistent with the 'gating by inhibition' hypothesis, which suggests that decreased alpha power translates to increased activity in the underlying cortex, and hence increased attentional resources (e.g., Jensen & Mazaheri, 2010).

For native speakers, we observed higher theta power for mismatching grammatical gender between a target noun and its preceding determiner compared to those where grammatical gender matched, while for L2 learners this was observed only when trials were sorted according to participants' subjective lexical representations, and only for *Experiment 2* where there were a relatively large number of trials (about twice as many as in *Experiments 3* and *4*). This suggests that these theta effects are relatively weak in L2 learners. There is clearly a relationship between induced oscillatory activity in the theta frequency range and language comprehension, but the precise nature of this link is not yet clear and warrants further research.

3.6.1 Beta oscillations and syntactic processing

Oscillatory activity in the beta frequency range has been linked to syntactic processing during sentence comprehension (see Lewis, Wang, & Bastiaansen, 2015 for review). On the other hand, not all types of syntactic manipulation modulate beta power (Davidson & Indefrey, 2007). Furthermore, the extent to which late second language learners show similar patterns of beta oscillatory activity for syntactic manipulations, and whether or not this is dependent on their subjective lexical representations, is not yet clear. In *Experiment 1* we showed that disrupting the processing of gender agreement between a noun and its preceding determiner modulates oscillatory power in the beta frequency range, while disruption of the processing of number agreement between determiner and noun does not. *Experiment 2* used the same stimuli to show that beta power in L2 learners is not modulated when gender agreement information is manipulated. *Experiment 3* ruled out that the absence of a beta effect for L2 learners was a result of confusion due to the repetition of nouns, once with the objectively correct and once with the objectively incorrect determiner. In *Experiment 4* we showed that when L2 learners are required to perform a task explicitly focusing their attention on grammatical information (a grammaticality/determiner judgment task), they exhibit a beta power modulation for disruptions of the processing of gender agreement information, comparable to that of the native speakers in *Experiment 1*.

In our study, L2 learners exhibited beta effects comparable to those for native speakers only when required to perform a grammaticality judgment task. The only other study investigating the link between oscillatory neural activity and sentence-level syntactic comprehension in L2 learners (Kielar et al., 2014) found no beta effects for L2 learners when participants were required to perform a grammaticality judgment task, but beta effects were present when these participants performed an acceptability judgment task. One important difference is that in the study by Kielar et al. (2014) there were semantic anomalies in addition to syntactic violations. This meant that in their grammaticality judgment task, participants had to respond to grammatically incorrect sentences, but to avoid responding to semantic anomalies. This requires inhibiting responses when semantic anomalies are present, which was not necessary in our grammaticality judgment task. In fact, the acceptability judgment task in Kielar et al. (2014), where participants simply had to judge whether or not a sentence was acceptable (pointing out syntactic, but also semantic violations), was more comparable to our grammaticality judgment task, and this likely explains the difference in findings. Task demands therefore appear to interact with the exact composition of the set of stimulus materials, and future research should explore this relationship further. A second important difference between L2 participants from Kielar et al. (2014) and those from our experiments is that their participants learned their second language far earlier in life (mostly before 12 years of age) than our L2 learners (most of whom learned Dutch between about 17 and 21 years of age). Their participants might therefore be better described as bilinguals (Kielar et al., 2014), since they report a very high level of proficiency in their second language (89.1 %; Kielar et al., 2014). This cannot be ruled out as an alternative explanation for the discrepancy between their results and those we report here.

Importantly, all sentences in *Experiment 4* were objectively syntactically correct, but when we sorted trials according to what our L2 participants considered syntactically correct and incorrect gender agreement relations between target nouns and their preceding determiners (presumably based on their subjective lexical representations), we observed similar beta effects to those observed for the native speakers (where subjective and objective lexical representations almost always overlap). This suggests that when L2 learners do use grammatical gender information during syntactic processing, it is their subjective lexical representation of that gender information, rather than the objectively correct information, that is most relevant. Our findings therefore corroborate the ERP findings of Lemhöfer et al. (2014) in arguing that for L2 learners who have not yet reached native-like proficiency, it is participants' idiosyncratic lexical-syntactic representations that count when it comes to syntactic processing.

3.6.2 Conclusions

These experiments show that grammatical gender agreement can be included amongst the now numerous factors related to syntactic processing that modulate oscillatory activity in the beta frequency range. Beta power is higher for nouns whose grammatical gender matches that of their preceding determiner, compared to those that exhibit a mismatch. Furthermore, late second language learners only exhibit similar beta oscillatory effects to native speakers when their attention is explicitly focused on grammatical information. These Beta effects in L2 learners are only observed when trials are sorted according to participants' idiosyncratic lexical representations of correct and incorrect gender agreement between target determiner-noun pairs. This suggests that L2 learners' subjective lexical representations matter more for syntactic processing than the objective correctness of the gender marking on the target noun. We also replicate the finding that grammatical number agreement is not one of the syntactic factors that modulates oscillatory activity in the beta frequency range, raising questions about the proposed link between beta and syntactic processing. Finally, theta power was also modulated by our syntactic manipulations for both native speakers and L2 learners, but these theta effects appear to be weaker for L2 learners, requiring a large number of trials to be observed. The exact nature of the relationship between induced oscillatory theta power and sentence-level language comprehension is not yet clear.

3.7 Epilogue

While the findings from the four experiments just described have been explicitly framed in terms of the 'beta-syntax' hypothesis, with lower beta power indicating disrupted syntactic processing, they are also perfectly compatible with the 'beta-maintenance' hypothesis. As I have argued in *Chapter 1*, under this hypothesis syntactic violations constitute cues to the language comprehension system, indicating that the current sentence-level meaning (and hence the underlying neural network configuration responsible for the construction and representation of that sentence-level meaning) needs to be revised. This is expected to result in lower beta power for syntactically incorrect compared to correct target words in a sentence. Native speakers, and L2 learners who are explicitly required to focus on grammatical information, exhibit exactly this pattern of results with higher beta power for syntactically correct compared to incorrect target words. Furthermore, if it is the case that German L2 learners of Dutch essentially ignore grammatical gender agreement information (adopting a shallow processing strategy; Ferreira & Patson, 2007) when they are not explicitly required to focus on this information (e.g., when reading for comprehension without any task requiring the use of grammatical information from the sentence), then the language comprehension system would not use violations of gender agreement as cues indicating a need to revise sentence-level meaning. This explains the absence of any differences in beta oscillatory activity for the Gender conditions in *Experiments 2* and *3*.

One problem for this interpretation is the following: if syntactic violations constitute cues to the language comprehension system indicating that the current mode of processing needs to change, why do neither native speakers (*Experiment 1*; Davidson & Indefrey, 2007) nor L2 learners (*Experiment 2*) exhibited beta effects for violations of number agreement? As alluded to earlier in the chapter, this anomaly is equally problematic for the 'beta-syntax' hypothesis. It is possible that beta effects are present for number agreement violations, but that a vastly improved signal-to-noise ratio (and hence a larger number of syntactically correct and incorrect trials) is necessary in order to detect any beta differences. Future research explicitly focusing on the relationship between grammatical number agreement processing and beta oscillatory dynamics is clearly warranted (here the Number condition was not our main focus).

Importantly, the findings from this chapter are equally supportive of both the 'beta-syntax' hypothesis and the 'beta-maintenance' hypothesis. They illustrate that these hypotheses apply equally well to the comprehension of late second language learners as they do to that of native speakers.

Notes

1. Dutch neuter and common gender nouns in singular form can also be preceded by the indefinite determiner een, which is not marked for gender. Our focus in this study is on gender agreement processing and so on definite determiners only.

2. Stimulus materials for *Experiments 1* and *2* were those used by Lemhöfer et al. (2014) and so our description of their construction follows that of Lemhöfer et al. (2014) very closely.

Chapter 4

Memory-related Reinstatement of Oscillatory Entrainment:

How Strong Are the Reinstatement Effects?

Abstract

Reinstatement of memory-related neural activation measured with high temporal precision offers a potentially very useful index of exactly when information from long-term memory becomes activated. An area where this could be particularly beneficial is in the tracking of lexical information as it is activated during language comprehension. Essential for such an approach is that the memory reinstatement effects are strong and robust, so that their absence can be taken as an indication that no lexical activation is present. In this study we used electroencephalography (EEG) to test the robustness of a reported subsequent memory finding involving the reinstatement of frequency-specific entrained oscillatory brain activity during recognition. Participants learned lists of words presented on a background, flickering at either 6 or 15 Hz to entrain the brain in a steady-state response, measured in the EEG signal. Later, participants saw the target words along with distractor words, this time presented on a background that did not flicker, and they had to indicate whether or not it was a word they had learned. For correctly recognized target words a reinstatement effect was found at 15 Hz for words encoded on a background that flickered at 15 Hz compared to words encoded on a background that flickered at 6 Hz. The reverse was not the case at 6 Hz. We thus partially replicated the original memory reinstatement findings, but conclude that these memory reinstatement effects are not robust enough to be used as a reliable index of lexical activation during language comprehension.

Keywords: EEG; memory reinstatement; steady-state brain response; frequency-specific oscillatory entrainment; subsequent memory; language comprehension; lexical activation; predictive processing

4.1 Introduction

Prediction and predictive processing has become a topic of great interest within cognitive science and cognitive neuroscience in recent years (e.g., Bubic, von Cramon, & Schubotz, 2010; Clark, 2013; Kveraga, Ghuman, & Bar, 2007). This is no less the case for language processing (e.g., Federmeier, 2007; Huettig, 2015; Pickering & Garrod, 2007). One key aspect of predictive processing during language comprehension is that it can lead to the activation of lexical information stored in long-term memory, even before that information appears in the linguistic input (e.g., DeLong, Urbach, & Kutas, 2005; Otten & Van Berkum, 2008; Szewczyk & Schriefers, 2013; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; Wicha, Moreno, & Kutas, 2004). Exactly how often predictive processing plays a role in language comprehension, and whether or not it forms an essential or fundamental part of the language comprehension system, are open questions. To address these issues, it would be useful to have a robust measure of exactly when during language comprehension different types of linguistic information become active.

One set of theories claim that for episodic memory, remembering entails the reinstatement of brain activity that was present when the memory was initially formed (e.g., Alvarez & Squire, 1994; Marr, 1971; McClelland, McNaughton, & O'Reilly, 1995; Norman & O'Reilly, 2003; Teyler & DiScenna, 1986). Neuroimaging work (predominantly using functional magnetic resonance imaging) has provided support for such 'memory reactivation' accounts of remembering (e.g., Johnson & Rugg, 2007; Johnson, McDuff, Rugg, & Norman, 2009; McDuff, Frankel, & Norman, 2009; Nyberg, Habib, McIntosh, & Tulving, 2000; Polyn, 2005; Vaidya, Zhao, Desmond, & Gabrieli, 2002; Wheeler, Petersen, & Buckner, 2000; Wheeler & Buckner, 2004; Woodruff, Johnson, Uncapher, & Rugg, 2005). These methods however offer relatively poor temporal resolution, which means they are not well suited to investigating the precise timing of memory reinstatement effects. A recent study has instead used electroencephalography (EEG), and the well-known phenomenon of entrainment of frequency-specific steady-state oscillatory brain responses, to address exactly this timing issue (Wimber, Maaß, Staudigl, Richardson-Klavehn, & Hanslmayr, 2012).

In a standard subsequent memory paradigm, Wimber et al. (2012) presented participants with a list of target words to be learned during an encoding phase of their experiment, and then after a short retention interval presented the same words along with new distracter words in a recognition phase. The key difference from previous studies was that during the encoding phase, words were presented on a background that flickered regularly at either 6 Hz or 10 Hz. Such flickering visual stimuli produce a steady-state visual evoked brain response in the EEG at the frequency of stimulation (Herrmann, 2001). These frequency-specific entrainment effects were hypothesized to be incorporated into the episodic memory representation participants formed during encoding, so that memories could be 'tagged' with one of the two stimulation frequencies. This allowed Wimber et al. (2012) to show early reinstatement of this frequency-specific brain activity during the recognition phase (when the background on which the words appeared did not flicker), such that words presented during encoding on a background that flickered at 6 Hz resulted in larger phase consistency in the EEG signals at 6 Hz than did words presented on a background that flickered at 10 Hz, and vice versa at 10 Hz.

Although episodic memory is not typically implicated in the core processes of language comprehension (e.g., Ullman, 2004), under the right set of circumstances (e.g., when a word is memorized shortly prior to reading a sentence, or sentences, containing that word) episodic information related to long-term memory representations for lexical items may become activated in addition to the lexical representations themselves (e.g., Glenberg et al., 2009; Tulving et al., 1994). We therefore proposed to use the approach from Wimber et al. (2012) to monitor when various lexical items become activated during sentence reading. By first 'tagging' target lexical items with a specific entrainment frequency during a memorization phase, we predict that it should be possible, in a rapid serial visual presentation paradigm, to monitor when (at which word of the sentence) these frequency tags become reinstated while people are reading each word of a sentence. If this reinstatement can be shown to be a robust indicator of memory (re)activation, then we would have a useful tool for (amongst a host of other things) monitoring whether lexical information is activated prior to the appearance of that information in the linguistic input.

Before applying this frequency tagging approach to sentence reading, we decided to replicate the reported finding to ensure that the frequency-specific reinstatement of brain activity that was present during encoding is a robust measure of memory reactivation during recognition. In the present study we therefore performed a close replication of the original study by Wimber et al. (2012). Participants were presented target words during an encoding phase of the experiment, in which they had to memorize the words and indicate whether the word was concrete or abstract. After a retention interval, participants were presented the target words along with new distracter words, and they had to indicate whether or not the word had appeared during the encoding phase.

During the encoding phase (but not during the recognition phase), the background on which the words were presented flickered regularly at either 6 or at 15 Hz (henceforth, the 6 Hz tagging condition and the 15 Hz tagging condition respectively). Participants' EEG was measured during both encoding and recognition phases of the experiment, allowing us to quantify the degree of frequency-specific phase consistency across trials in the EEG signal, by calculating inter-trial coherence (ITC; Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996). The difference in ITC at both 6 Hz and at 15 Hz between trials from the 6 Hz tagging condition and trials from the 15 Hz tagging condition was used to quantify whether a steady-state evoked response was present (during the encoding phase), and whether frequency-specific memory reinstatement occurred (during the recognition phase).

We hypothesized that if the steady-state brain responses during encoding become associated with the memory representations formed for the memorized words, then during the recognition phase we should observe greater phase consistency at 6 Hz for trials from the 6 Hz tagging condition compared to trials from the 15 Hz tagging condition, and vice versa for phase consistency at 15 Hz.

4.2 Methods

The methods closely followed those employed by Wimber et al. (2012).

4.2.1 Participants

Twenty-eight native speakers of Dutch took part in the experiment, 21 of whom were included in the final analysis (8 males, 13 females; aged 19 to 36). Participants provided informed consent and were paid or equivalently rewarded with course credits for their participation. All participants reported normal or corrected-to-normal vision, and were right handed. None of the participants reported any neurological impairment.

Three participants were excluded from the final analysis due to recording problems. Four other participants were excluded due to poor data quality, which resulted in too few trials for the ITC analysis (less than 35 hit trials in either of the frequency tagging conditions during the recognition period).

4.2.2 Stimulus materials

Target stimuli consisted of 200 Dutch words selected from a large database of Dutch-English translation pairs (Tokowicz, Kroll, de Groot, & van Hell, 2002)[1]. Approximately two thirds of the words (67.5 %) referred to concrete entities (e.g., *table* or *carrot*), while a third (32.5 %) referred to abstract concepts (e.g., *revenge* or *wisdom*).

A total of 100 Dutch words were selected as distracter items to be used during the recognition phase of the experiment. Every participant saw all 200 target and 100 distracter stimuli over the course of the experiment. Which target and distracter items were seen in the first and second blocks of the experiment, as well as which half of the distracter items was presented together with which half of the target items in any single block, was counterbalanced across participants (for a description of the blocks see *Experimental design and procedure* below). Whether a target word was presented on a background flickering at 6 or at 15 Hz was also counterbalanced across participants. Resulting experimental lists for each block were then pseudorandomized according to the following criteria (the same for both encoding and recognition phase experimental lists): 1) no more than three consecutive repetitions of concrete or abstract words; 2) words from the same semantic category appeared at least three experimental items apart; 3) no more than two consecutive repetitions of a word presented with the same frequency of flickering background during the encoding phase (6 Hz or 15 Hz).

During the retention interval participants performed a so-called flanker task (see below). Stimuli for this task consisted of a central arrow pointing either to the left or to the right of the display. These were 'flanked' to both the left and the right by either three arrows or three equals signs (neutral condition). The flanker arrows could either point in the same direction as the central arrow (congruent condition) or in the opposite direction to the central arrow (incongruent condition). This task was only used to ensure that participants did not have time for rehearsal of the learned words during the retention interval, and thus these data were not analysed.

4.2.3 Experimental design and procedure

Participants were tested in a dimly lit, sound-attenuating and electrically shielded booth. They were seated comfortably in front of an LCD computer monitor, with a viewing distance of between 70 and 80 cm. Words were presented in white on a rectangular black box using a 20-point sized uppercase Consolas font type. All words subtended a visual angle of 3.13° vertically. The black box measured 350x300 dpi, and was presented centrally on a dark grey background encompassing the remainder of the display (1920x1080 dpi).

The experiment consisted of two blocks, each comprised of three experimental phases. First, in an encoding phase participants learned 100 Dutch words while judging whether the words presented were concrete or abstract. We used a so-called 'deep encoding' task rather than the 'shallow encoding' task used in Wimber et al. (2012). In principle this should not make any difference to the memory reinstatement effects, and since we planned to use the present data as a basis for follow up experiments targeting the semantics of the words, focusing participants' attention on semantics by using a deep encoding task seemed most appropriate. Directly following the encoding phase participants completed a flanker task (90 trials: 30 congruent; 30 incongruent; 30 neutral) that was unrelated to our experimental manipulation. This was used to ensure that participants did not have the opportunity to rehearse learned words during this 5- to 6-minute retention interval. Finally, in the recognition phase of the experiment participants saw target words they had learned during the encoding phase for that block, interspersed with 50 new Dutch distracter words they had not seen during encoding. They were required to judge whether or not these were words they had seen during encoding, and to indicate how certain they were of their response. Before the start of the first block participants completed a short practice session with items not used in the main experiment in order to ensure that they understood the task for each phase. Each block lasted approximately 40 minutes, with the entire experiment (including preparation and practice) taking around 2 hours.

During the encoding phase of the experiment, each trial began with the presentation in the centre of the screen of three asterisks two spaces apart for 2000 ms, indicating that participants could move their eyes and blink. This was immediately followed by a fixation cross presented in the centre of the screen for between 1000 and 1750 ms, indicating that eye movements and blinking should be avoided and the word was about to appear. Next, a word was presented in the centre of the screen for 2500 ms. During the presentation of the word the black background box flickered regularly at either 6 or 15 Hz, changing from black (RGB: 0, 0, 0) to the dark grey (RGB: 125, 125, 125) colour of the display background and back to black. Directly following the word, a question mark was presented in the centre of the screen for a maximum duration of 1500 ms, indicating that participants should provide a (subjective) judgment about whether the word referred to something abstract or concrete. The question mark disappeared immediately upon a response from the participant, and the next trial began. Participants were instructed to remember the words as they would be tested in the recognition phase of the experiment on whether or not they had seen them during the encoding phase. The subjective judgment about whether a word referred to something concrete or abstract was made by a button press with either their left or right index finger. For half the participants the left index finger was used for words they thought were concrete and the right for words they thought were abstract, and vice versa for the other half of the participants.

During the recognition phase of the experiment, trials were the same as during the encoding phase, except that this time words were presented for 2000 ms and the black background box did not flicker. Participants were instructed to press a button with one hand in case they had seen the word during the encoding phase of the current block, or with the other hand in case it was a word they had not seen during the encoding phase. For each of these options participants had to indicate how certain they were about their response, with the index finger indicating complete certainty, the middle finger indicating medium certainty, and the ring finger indicating complete uncertainty. Response mapping was again counterbalanced across participants.

During the retention interval each trial began with the presentation in the centre of the screen of three asterisks two spaces apart for between 1500 and 2500 ms, indicating that participants could move their eyes and blink. This was immediately followed by a fixation cross presented in the centre of the screen for 500 ms, indicating that eye movements and blinking should be avoided and the stimulus was about to appear. Next, the flanker stimuli (without the central target arrow) were presented in the centre of the screen for 100 ms, directly followed by the flanker stimuli with the central target arrow for 50 ms. Finally, a question mark was presented in the centre of the screen for a maximum duration of 1500 ms, indicating that participants should use their left or right index finger to press the button corresponding to the direction in which the central arrow was pointing. The question mark disappeared immediately upon a response from the participant, and the next trial began.

4.2.4 EEG recordings

Participants were fitted with a 59 electrode cap with electrodes positioned in the geodesic arrangement shown in Figure 4.1. EEG signals were recorded using 59 Ag/AgCl active sensors mounted in the cap and referred to the left mastoid. An additional electrode was placed on participants' right mastoid, and a ground electrode was placed on the centre of the forehead. Another electrode was placed on the suborbital ridge of participants' left eye for recording eye-blinks.

Electrode impedance was kept below 20 kΩ. EEG and EOG recordings were amplified using BrainAmp DC amplifiers (Brain Products GmbH, Gilching, Germany) with a band-pass filter of 0.016 to 200 Hz, digitized online with a sampling frequency of 500 Hz and stored for offline analysis.

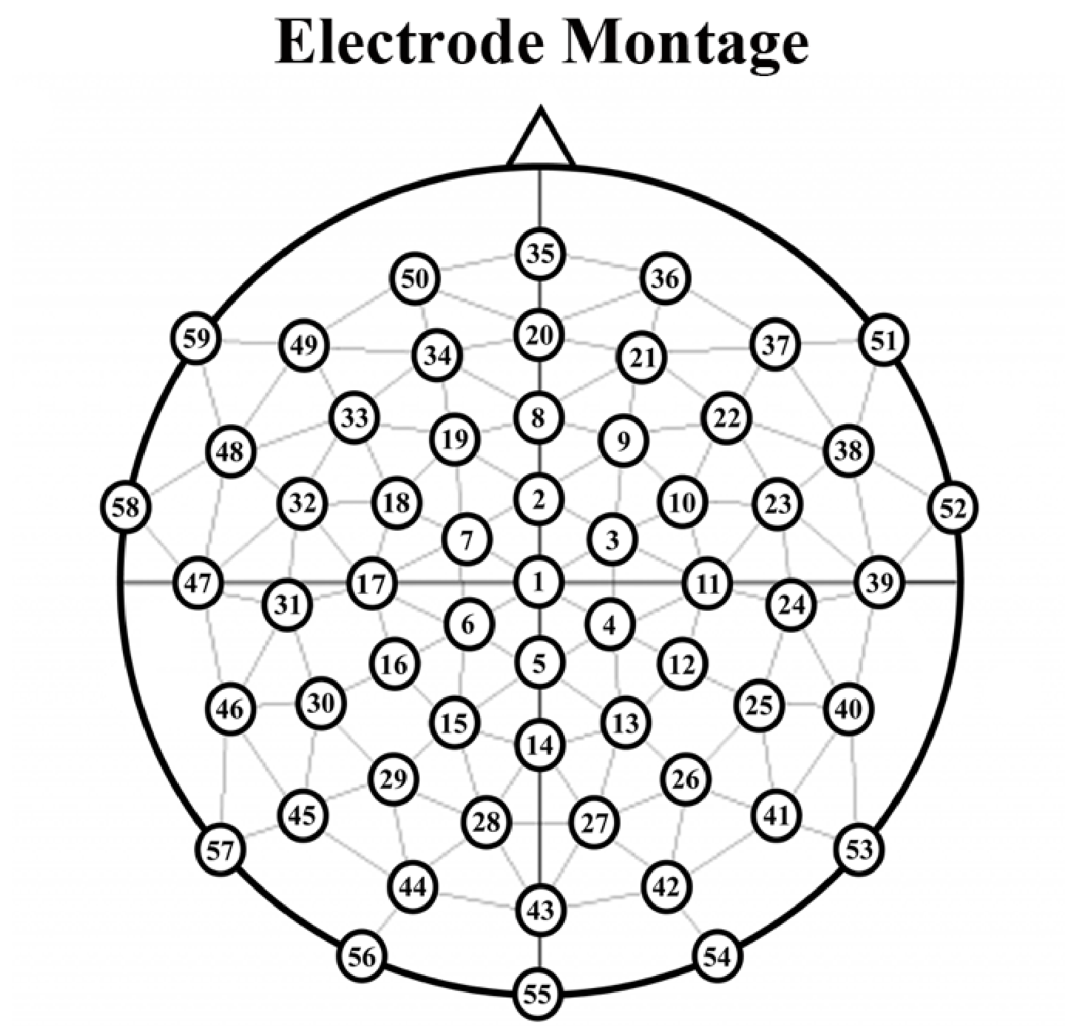

Figure 4.1 Positions of the scalp electrodes in the EEG cap.

4.2.5 Behavioural data analysis

Recognition performance was assessed separately for the 6 Hz and 15 Hz tagging conditions, using the signal detection measures recognition sensitivity (d′) and response bias (β), as well as by plotting individual (per participant) and mean receiver operating characteristic (ROC) curves (e.g., Stanislaw & Todorov, 1999; Yonelinas & Parks, 2007). For plotting the ROC curves, participants' responses during the recognition phase of the experiment were ordered according to the level of certainty of their response from C1 (completely certain seen word during encoding phase) to C6 (completely certain haven't seen word during encoding phase). The hit rate was calculated as the proportion of target words from the encoding phase correctly identified as previously seen, while the false alarm rate was calculated as the proportion of distracter words incorrectly identified as seen during the encoding phase. These scores were used to calculate recognition sensitivity and response bias for each participant. A paired-samples T-test was then used to assess whether there were any systematic differences between the 6 Hz and 15 Hz tagging conditions in either of these measures across participants. Trials in which participants failed to make a response were excluded from the above analyses. This behavioural analysis is based only on data from participants who were included in the final EEG analysis.

4.2.6 EEG data pre-processing

EEG data were analysed using the FieldTrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011) running in a MatLab environment (R2014b; Mathworks, Inc.). The following pre-processing steps were applied separately to the data from the encoding and the recognition phases of the experiment. For each participant, a band-stop filter was applied at 50, 100, and 150 Hz in order to minimize the effects of power line interference (50 Hz), and data were segmented from -1000 to 2500 ms relative to the onset of each word. The data were then visually inspected, and any electrodes exhibiting non-stationary artifacts in a large number of trials were removed from the data. Each electrode was then re-referenced to the average of all scalp electrodes (common average reference).

Next the data were decomposed into independent components (ICA using the 'runica' implementation in FieldTrip with default settings). Components which captured eye-blinks or horizontal and vertical eye movements were removed and the remaining components were recombined (Jung et al., 2000; Makeig, Jung, Bell, Ghahremani, & Sejnowski, 1997). Between 0 and 5 components were removed per participant. Electrodes that had been removed were then recovered based on the average activity at neighbouring electrodes, and a low-pass filter was applied at 30 Hz. Each data segment was then demeaned using the mean over the entire segment and any linear trends were removed. Any remaining artifactual trials were removed after visual inspection of all data segments.

Finally, the data were segmented into 6 Hz and 15 Hz tagging conditions from -500 to 2500 ms relative to word onset. For the data from the recognition phase of the experiment, only data segments corresponding to hit trials (regardless of participants' certainty rating) were included for further analysis. There were no statistically significant differences between the number of remaining trials in the two conditions for either the encoding (6Hz: $M = 61.95$, SD = 7.66; 15 Hz: $M = 64$, SD = 8.29; p = 0.24) or the recognition (6 Hz hits: $M = 56.81$, SD = 11.16; 15 Hz hits: $M = 57.43$, SD = 8.93; p = 0.74) phase of the experiment.

4.2.7 Inter-trial coherence analysis

Our analysis of the EEG data focused on the degree of phase consistency across trials relative to stimulus onset within a given frequency band. To that end, we computed the inter-trial coherence (ITC; Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996; sometimes referred to as the phase-locking index; e.g., Wimber et al., 2012) for each participant from 0 ms to 2000 ms relative to the onset of each target word. For all trials of the encoding phase and hit trials of the recognition phase of the experiment, ITC was computed separately for the 6 Hz and 15 Hz tagging conditions. First, time-resolved Fourier spectra of the data between 2 and 22 Hz were computed using a sliding window approach. Sliding windows of 1000 ms were applied in frequency steps of 1 Hz and time steps of 20 ms, and each window was tapered using a Hanning taper to reduce spectral leakage. This resulted in a frequency resolution of 1 Hz, while the estimate at each time point is averaged data from the preceding and following 500 ms. Next, the Fourier spectrum of each trial was normalized by its amplitude, and the result was averaged across all trials for a particular tagging condition. This provided a frequency-resolved measure of the degree of trial-to-trial phase consistency over time (ITC).

4.2.8 Statistical analyses

For data from both the encoding and recognition phases of the experiment, statistical significance was evaluated using the approach described by Wimber et al. (2012). In a first step, non-parametric Wilcoxon signed-rank tests were used to compare the 6 Hz and 15 Hz tagging conditions for every time-frequency pair at each electrode ($p < 0.05$ considered significant; one-tailed test). This resulted in uncorrected P-values for every time-frequency pair at each electrode. A time window of interest was then selected for data points at 6 Hz and at 15 Hz by selecting the largest time window exhibiting adjacent significant time points on at least 4 electrodes. In a second step, a randomization approach was used to ensure that the number of electrodes showing a significant difference between the average ITC over the time window of interest in the 6 Hz (positive one-tailed test) and 15 Hz (negative one-tailed test) frequency range was higher than would be expected by chance ($p_{corr} < 0.05$ considered significant; see Blair & Karniski, 1993; Hanslmayr, Spitzer, & Bäuml, 2009 for more extensive descriptions of the approach). For the recognition phase only hit trials from each tagging condition were included in the statistical analyses.

In addition to the statistical testing used by Wimber et al. (2012) we carried out cluster-based random permutation tests (Maris & Oostenveld, 2007) in order to compare these with the main statistical results, and to investigate how robust the findings are. In short, a non-parametric Wilcoxon signed-rank test was performed for every data point (electrode-time point) in the 6 Hz frequency range (positive single-tailed test) and in the 15 Hz frequency range (negative single-tailed test) separately, comparing trials from the 6 Hz and 15 Hz tagging conditions. A pre-set significance level was chosen (here 5% single-tailed) and any data points not exceeding this level were discarded (set to zero). Clusters were calculated from the remaining data points based on their adjacency in space (adjacent electrodes) and time.

Cluster-level statistics were then calculated by summing the resultant Z-values for all data points in each cluster. A permutation distribution was created by randomly assigning participant averages to one of the two conditions 3000 times, and each time calculating cluster-level statistics as just described. The highest cluster-level statistic from each randomization was entered into the permutation distribution and the cluster-level statistics calculated for the measured data were compared against this distribution (cluster corrected $p < 0.05$ considered significant).

4.3 Results

4.3.1 Behavioural Results

There were no statistically significant differences in recognition performance between the 6 Hz (d'_{6Hz}: *M* = 1.99, SD = 0.64; β_{6Hz}: *M* = 0.59, SD = 0.24) and the 15 Hz (d'_{6Hz}: *M* = 1.95, SD = 0.55; β_{6Hz}: *M* = 0.63, SD = 0.26) tagging conditions (d′: t_{20} = 0.54, p = 0.6; β: t_{20} = -0.82, p = 0.42). Figure 4.2 shows overlapping ROC curves for the two tagging conditions. These results suggest that the two frequencies of flickering background did not differentially affect the encoding of the target words. For both tagging conditions, the d′ results indicate that participants were clearly able to distinguish target from distractor words, and that there was a bias (β) toward responding that a word had already been seen during encoding. The latter finding is not surprising given that during recognition two thirds of the items were target words and only one third were distractor words.

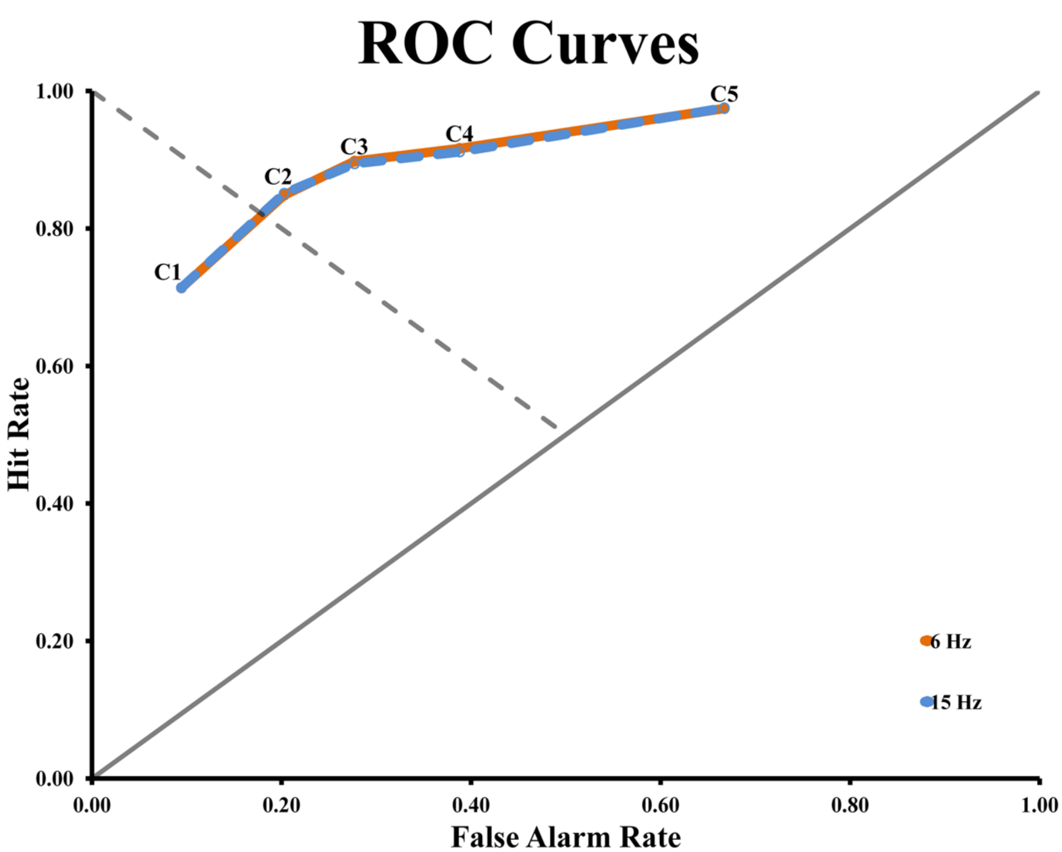

Figure 4.2 Receiver operating characteristic (ROC) curves for the 6 Hz and 15 Hz tagging conditions. Cumulative hit rate plotted against cumulative false alarm rate for different levels of confidence (C1-C6). Solid diagonal line = chance performance; dashed diagonal line = neutral response criterion.

4.3.2 ITC Results

ITC values for the items from the two tagging conditions were compared in a 6 Hz (6 Hz tagging condition > 15 Hz tagging condition) and in a 15 Hz (15 Hz tagging condition > 6 Hz tagging condition) frequency range.

4.3.2.1 Encoding phase

Based on the statistical procedure used by Wimber, et al. (2012), between 340 and 1900 ms relative to word onset, words from the 6 Hz tagging condition exhibited stronger phase locking at 6 Hz than words from the 15 Hz tagging condition ($p_{corr} < 0.001$). Similarly, between 320 and 1900 ms relative to word onset, words from the 15 Hz tagging condition exhibited stronger phase locking at 15 Hz than words from the 6 Hz tagging condition ($p_{corr} < 0.001$). Figures 4.3A and 4.3B show ITC differences and uncorrected P-values averaged over all electrodes showing a significant difference across the entire time interval of interest (after correcting for multiple comparisons), between the two tagging conditions at 6 Hz (all 59 electrodes; Figure 3A) and at 15 Hz (51 out of the 59 electrodes; Figure 4.3B).

The cluster-based statistical approach yielded very similar results. Between 300 and 1900 ms relative to word onset, words from the 6 Hz tagging condition exhibited stronger phase locking at 6 Hz than words from the 15 Hz tagging condition ($p_{corr} < 0.001$; significant at 59 electrodes). Similarly, between 280 and 1900 ms relative to word onset, words from the 15 Hz tagging condition exhibited stronger phase locking at 15 Hz than words from the 6 Hz tagging condition ($p_{corr} < 0.001$; significant at 58 electrodes).

These results confirm that we were able to entrain a steady state brain response at the frequency corresponding to the frequency of the flickering background during the encoding phase of the experiment. Furthermore, both statistical approaches employed confirmed the effect.

4.3.2.2 Recognition phase

Only trials where target words were correctly recognized as having been presented during the encoding phase of the experiment (hit trials) were included in the statistical analyses. Based on the statistical procedure used by Wimber, et al. (2012), between 780 and 1000 ms relative to word onset, words from the 15 Hz tagging condition exhibited stronger phase locking at 15 Hz than words from the 6 Hz tagging condition (significant at 12 electrodes; $p_{corr} = 0.003$). There was no

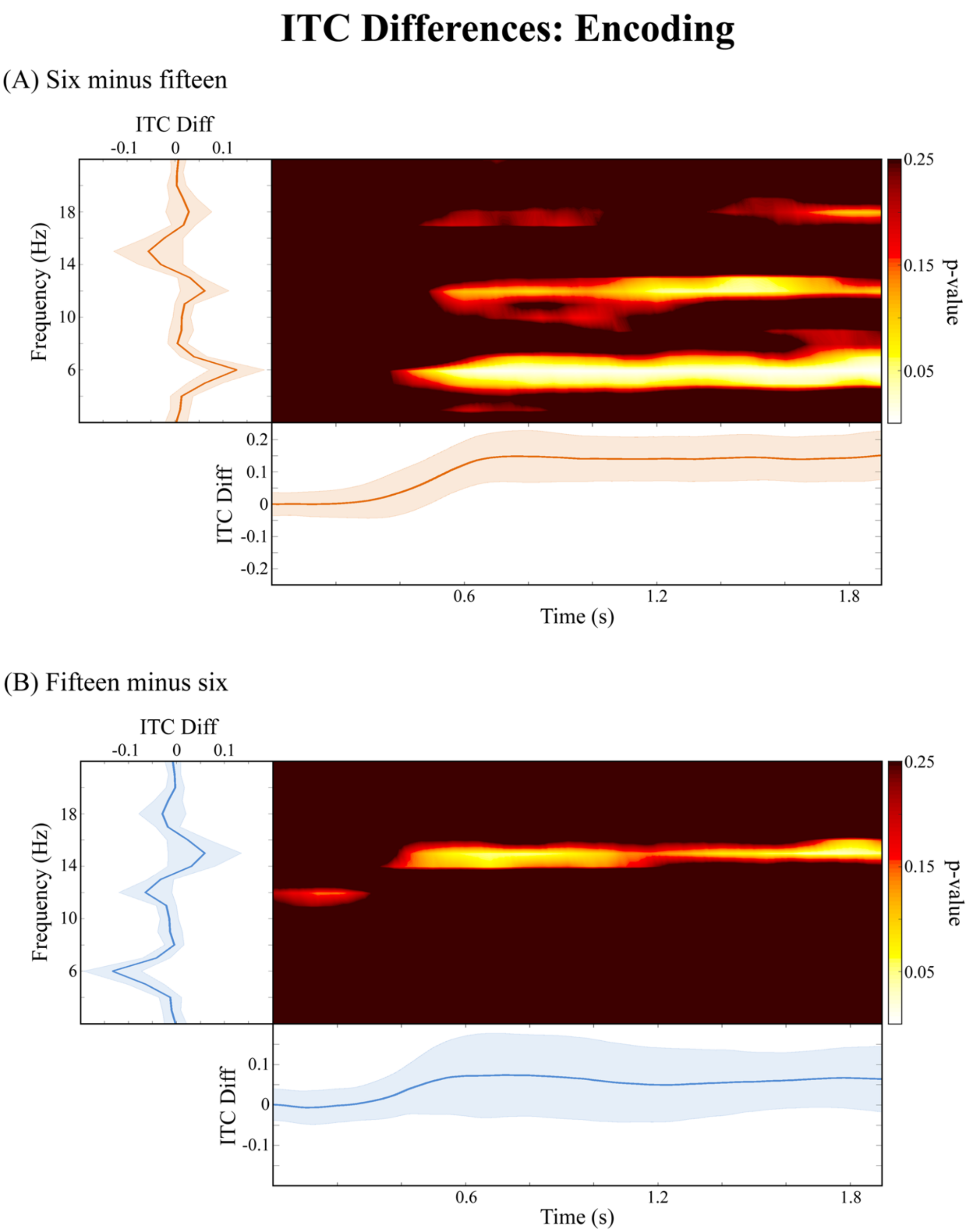

Figure 4.3 ITC differences during encoding. (A) ITC difference between target words from the 6 Hz tagging condition compared to target words from the 15 Hz tagging condition. Significantly more phase consistency was present at 6 Hz for words from the 6 Hz tagging condition than for words from the 15 Hz tagging condition (340 to 1900 ms; 59 electrodes). Differences are also present at harmonic frequencies (12 Hz and 18 Hz). (B) ITC difference between target words from the 15 Hz tagging condition compared to target words from the 6 Hz tagging condition. Significantly more phase consistency was present at 15 Hz for words from the 15 Hz tagging condition than for words from the 6 Hz tagging condition (320 to 1900 ms;

51 electrodes). Time-frequency plots show uncorrected P-values averaged over all electrodes showing a significant entrainment effect; line plots to the left show average ITC differences over the significant time interval and significant electrodes as a function of frequency; line plots at the bottom show average ITC differences over significant electrodes at the frequency of interest as a function of time; shaded regions indicate standard error of the mean.

statistically significant difference in phase locking at 6 Hz between words from the 6 Hz tagging condition and words from the 15 Hz tagging condition ($p_{corr} = 0.11$). Figures 4.4A and 4.4B show ITC differences and uncorrected P-values averaged over all electrodes. At 15 Hz there is a significant difference across the entire time interval of interest (after correcting for multiple comparisons) between the two tagging conditions (Figure 4.4B). At 6 Hz there is a difference between the two tagging conditions between 500 and 600 ms relative to word onset (at 6 electrodes) which however does not reach statistical significance (Figure 4.4A).

The cluster-based statistical approach yielded no statistically significant differences between the two tagging conditions at either 6 Hz ($p_{corr} = 0.92$) or at 15 Hz ($p_{corr} = 0.11$). We thus replicate the memory reinstatement effect observed by Wimber et al. (2012), but only for one of our two frequency tagging conditions (15 Hz tagging condition), and only when using the statistical approach employed by Wimber et al. (2012). When employing an alternative statistical technique for comparing the two tagging conditions no statistically significant differences are found.

4.4 Discussion

Memory-related reinstatement of frequency-specific EEG activity, measured with high temporal precision, offers a potentially very useful method for tracking exactly when information stored in long-term memory becomes activated (possibly even before any external retrieval cue is present). One area where such a method could be particularly beneficial is in measuring lexical activation (or indeed pre-activation due to predictive processing) during language comprehension. Before employing such a ‘frequency tagging’ approach in a new context, we decided to first replicate the original finding (Wimber et al., 2012). Participants learned lists of target words during an encoding phase and later during a recognition phase saw lists of target and distractor words, and had to indicate whether or not they had seen each word during encoding. Target words were presented on a background that flickered at either 6 or 15 Hz during encoding to entrain a steady-state brain response in the EEG signal at one of these two flicker frequencies.

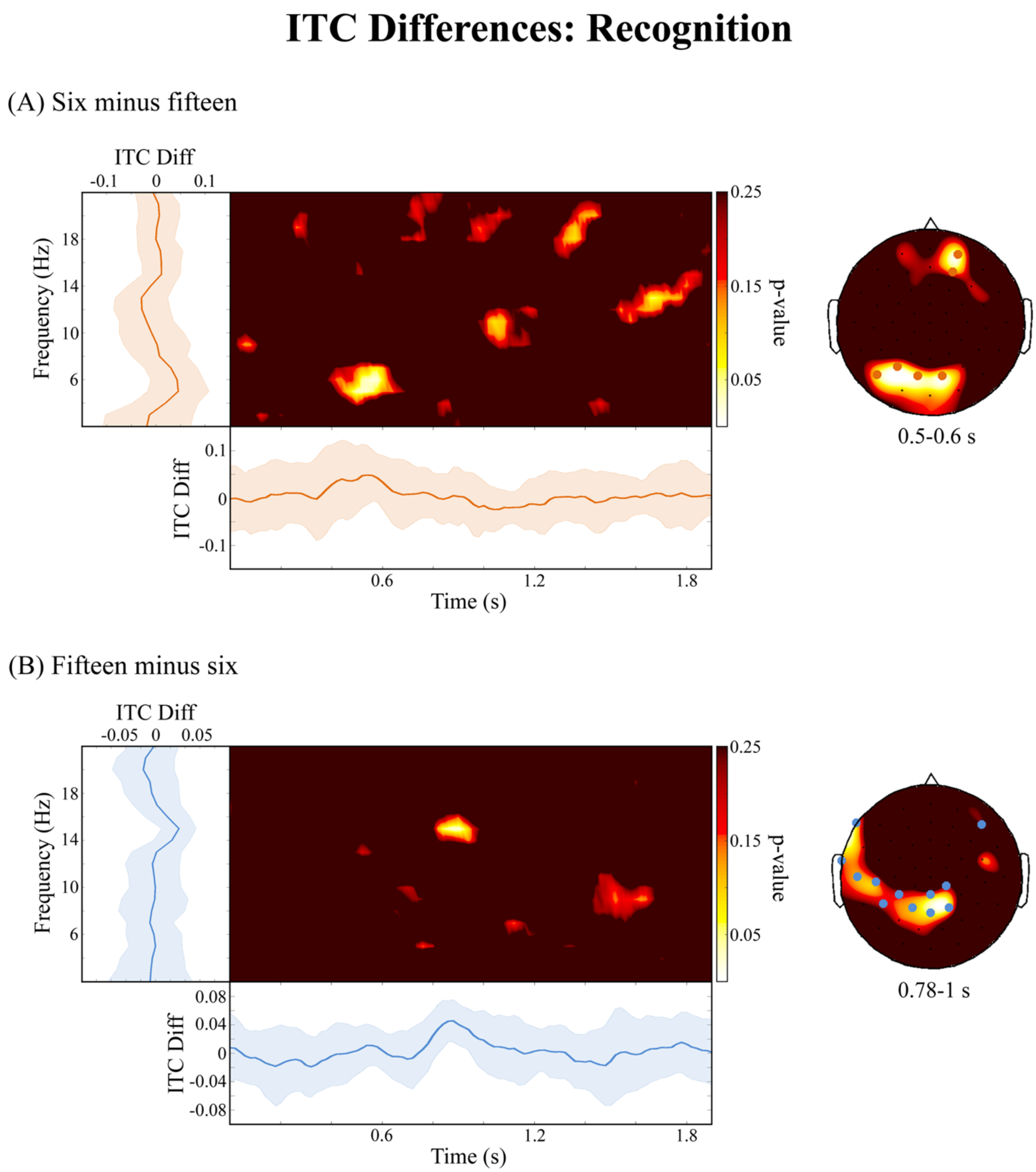

Figure 4.4 ITC differences during recognition. (A) ITC difference between correctly recognized target words from the 6 Hz tagging condition compared to correctly recognized target words from the 15 Hz tagging condition. More phase consistency was present at 6 Hz for words from the 6 Hz tagging condition than for words from the 15 Hz tagging condition, but this difference was not significant after correcting for multiple comparisons (500 to 600 ms; 6 electrodes). (B) ITC difference between correctly recognized target words from the 15 Hz tagging condition compared to correctly recognized target words from the 6 Hz tagging condition. Significantly more phase consistency was present at 15 Hz for words from the 15 Hz

tagging condition than for words from the 6 Hz tagging condition (780 to 1000 ms; 12 electrodes). Time-frequency plots show uncorrected P-values averaged over all electrodes showing a significant reinstatement effect (B) or over all electrodes showing a significant difference (uncorrected) for the entire time interval of interest (A); line plots to the left show average ITC differences over the time interval of interest and over all electrodes showing a significant difference (uncorrected) for the entire time interval of interest as a function of frequency; line plots at the bottom show average ITC differences over all electrodes showing a significant difference (uncorrected) for the entire time interval of interest at the frequency of interest as a function of time; shaded regions indicate standard error of the mean; scalp plots to the right show uncorrected P-values averaged over the time interval of interest for the frequency of interest; electrodes showing difference over entire time interval of interest are marked (orange = 6 Hz comparison, blue = 15 Hz comparison).

4.4.1 Frequency-specific oscillatory entrainment during encoding

The ITC analysis yielded the expected steady-state response during encoding, with greater phase consistency at 6 Hz for target words from the 6 Hz tagging condition than for those from the 15 Hz tagging condition, and greater phase consistency at 15 Hz for target words from the 15 Hz tagging condition than for those from the 6 Hz tagging condition (Figure 4.3). These effects were confirmed by both statistical procedures employed, and like in the original study by Wimber et al. (2012), were significant at almost all scalp electrodes (all 59 for the 6 Hz tagging condition and 51 out of 59 for the 15 Hz tagging condition).

One difference from the original study is that our steady-state entrainment effects at both 6 and 15 Hz appear to start around 300 ms after the onset of the target word. Wimber et al. (2012) found entrainment effects starting from the onset of the target word. This may simply be due to differences in temporal smoothing for the ITC calculation between our study and the study by Wimber et al. (2002). On the other hand, one might expect that it takes some time after the onset of a flickering stimulus before a steady state oscillatory response builds up. In any case, most important is that both studies show the presence of a frequency-specific steady-state oscillatory response at the frequency of flicker, and this could lead to the formation of an association between the oscillatory brain activity and the memory representation formed during encoding.

4.4.2 Memory-related reinstatement during recognition

Our ITC analysis partially replicated the memory reinstatement findings from Wimber et al. (2012). Higher phase consistency was found at 15 Hz for hit trials with target words from the 15 Hz tagging condition than for hit trials with target words from the 6 Hz tagging condition. This memory reinstatement effect was present between 780 and 1000 ms relative to target word onset.

There was however no statistically significant difference in phase consistency at 6 Hz (Figure 4.4). Furthermore, the effect at 15 Hz was only confirmed using the original statistical approach from Wimber et al. (2012), but not when a cluster-based permutation approach was employed (Maris & Oostenveld, 2007). The absence of a statistically significant memory reinstatement effect at 6 Hz may be due to the large degree of variability in the data (shaded region in Figure 4.4A showing the standard error of the mean), as visual inspection indicates at least a hint of a memory reinstatement effect between 500 and 600 ms after target word onset.

The first thing to notice is that our memory reinstatement effect at 15 Hz occurs later (starting around 780 ms after target word onset) than the effects reported by Wimber et al. (2012; memory reinstatement in their 6 and 10 Hz tagging conditions occurred between 0 and 300 ms relative to target word onset). This is also the case for our largest time interval showing a difference (not statistically significant) in phase consistency between the 6 and 15 Hz tagging conditions at 6 Hz (between 500 and 600 ms relative to target word onset). Like for the encoding phase, these differences may simply be due to the different smoothing parameters used in the two studies for the ITC calculation. Another possibility is that because we employed a deep encoding task (rather than the shallow encoding task employed by Wimber et al., 2012) the 'frequency tags' became associated with a different aspect of the memory representation formed when the target words were encoded into long term memory (possibly the semantics rather than just the word form). This has important implications for whether or not these memory reinstatement effects are indeed, as was argued by Wimber et al. (2012), an index of early ecphoric processing during memory retrieval (e.g., Moscovitch, 2008; Tulving, Voi, Routh, & Loftus, 1983). Further investigation is clearly warranted.

As was the case for Wimber et al. (2012), our reinstatement effect during recognition was present at a subset of the electrodes that exhibited an entrainment effect during the encoding phase, although since entrainment was observed at almost all scalp electrodes during encoding this is not at all surprising. More interesting is our finding that the 15 Hz effect and our non-significant difference at 6 Hz are present over largely spatially contiguous electrodes. Wimber et al. (2012) report their effect at 6 Hz over spatially non-contiguous electrodes (see their Figure 4.2C), with only their 10 Hz reinstatement effect showing spatial contiguity over left fronto-central electrodes. While the spatial distribution of memory reinstatement effects is likely related to the network of regions that were active during encoding (which could be widely distributed and are probably dependent on the nature of the stimulus set being encoded into memory) we think that there should still be a large degree of spatial contiguity when observing EEG activity at the scalp related to these effects.

Our 15 Hz reinstatement effect is present mainly over left temporal and over centro-parietal electrodes, while our 6 Hz non-significant difference is present more posteriorly over parieto-occipital electrodes (Figure 4.4). The scalp distribution and timing of these memory reinstatement effects raised the suspicion that they are related to the classical parietal old/new ERP memory effects (Rugg & Curran, 2007). On the other hand, our behavioural results indicate that there were no differences in memory performance between the 6 Hz and the 15 Hz tagging conditions, and we therefore think it unlikely that this could explain our findings.

4.4.3 Utility of memory-related reinstatement effects

Two further questions arise about the utility of these memory-related reinstatement effects. The first concerns whether or not reinstatement is functionally related to memory performance. Wimber et al. (2012) address this issue by performing two analyses comparing phase information for correctly recognized target word trials (hits trials) with phase information for trials where a target word was incorrectly judged as not seen during encoding (miss trials). We were not able to perform similar analyses with our data due to a low number of miss trials, which is not sufficient for obtaining a satisfactorily unbiased estimate of the phase information for each participant. We were therefore unable to test whether our 15 Hz memory reinstatement effect had any influence on behavioural memory performance. This was however never the main goal of the present study.

The second question does address the main goal of our study, and relates to whether or not a 'frequency tagging' and subsequent memory reinstatement approach could be used to investigate the timing of activation of lexical information during language comprehension. One prerequisite is that the memory reinstatement effects are strong and robust so that their presence, and more importantly their absence, can be taken as an indication of whether or not lexical information is activated. This is not the case in our study, as we show memory reinstatement only at 15 Hz but not at 6 Hz. Furthermore, the statistical significance of the memory reinstatement effect does not appear to be stable across changes in the preferred approach that is used to correct for the multiple statistical tests performed. The cluster-based random permutation approach (Maris & Oostenveld, 2007) we employed has proven a useful procedure for controlling the familywise error rate based on physiologically plausible assumptions. While it may be true that different statistical procedures are potentially sensitive to different aspects of the data (e.g., Groppe, Urbach, & Kutas, 2011), the fact that a Wilcoxon signed-rank test was used for both statistical approaches, along with the fact that the effect at 15 Hz appears to form a spatially and temporally contiguous cluster (which should work in favour of finding an effect with the cluster-based random permutation approach), raises the question whether the procedure employed by Wimber et al. (2012) is adequate for controlling for the multiple statistical tests performed. In our opinion this issue warrants further simulation studies to test whether or not the procedure may lead to spurious findings.

It is clear that the frequency-specific memory reinstatement effects we proposed to use for tracking lexical information during language comprehension are unfortunately not sufficiently strong and robust to be used as a reliable marker of lexical activation. While we cannot rule out that some of the minor changes we made to our experiment (deep rather than shallow encoding task; 15 Hz rather than 10 Hz frequency entrainment during encoding) cause our memory reinstatement effects to be less robust than those reported by Wimber et al. (2012), we do think that further investigation is needed before these effects can be used to track the activation of lexical representations during language comprehension.

4.4.4 Conclusions

This study partially replicates previous memory reinstatement findings (Wimber et al., 2012). We were able to entrain a frequency-specific steady-state brain response in the EEG signal at 6 and 15 Hz while target words were encoded into long-term memory. Later, during the recognition phase of the experiment, we observed higher phase consistency at 15 Hz for correctly recognized target words that were encoded on a background flickering at 15 Hz compared to target words that were encoded on a background flickering at 6 Hz. A similar memory reinstatement effect was not present at 6 Hz when comparing target words encoded on a background that flickered at 6 Hz to target words encoded on a background that flickered at 15 Hz. We conclude that these frequency-specific memory reinstatement effects are not sufficiently strong and robust to be used as a reliable marker of lexical activation during language comprehension.

Notes

1. Words were selected from this database because we planned to also use the English translations

of the Dutch words in future experiments.

Chapter 5

Is Beta in Agreement with the Relatives? Using Relative Clause Sentences to Investigate the Role of Beta Oscillations During Language Comprehension

Abstract

In this study we used magnetoencephalography to investigate the role of oscillatory activity in the beta frequency range during sentence-level language comprehension. Participants read relative clause sentences that were ambiguous between a subject- or an object-relative reading up to the point of disambiguation at a relative clause-final auxiliary. A third condition comprised violations of grammatical person agreement between the clause final auxiliary and either of the two preceding noun phrases. We were primarily interested in the modulation of beta power in response to the object-relative clause sentences, where an unexpected (or less expected) event occurs at the point of disambiguation, but unlike for agreement violation sentences the sentence remains grammatical. In an event-related field analysis we observed magnetic P600 effects for the comparison between agreement violation and subject-relative clause sentences, as well as for the comparison between object-relative and subject-relative clause sentences. Beta power over left frontal and left temporal sensors decreased for both agreement violation and for object-relative clause sentences, but showed no such decrease for subject-relative clause sentences. This provides compelling support for the idea that during sentence comprehension, beta activity is related to the maintenance/change of the network configuration responsible for the representation and construction of the current sentence-level meaning, and against the idea that such beta activity is related to syntactic unification operations.

Keywords: relative clause; MEG; beta oscillations; subject-relative; object-relative; agreement violations; ERF; P600

5.1 Introduction

One of the most striking aspects of language comprehension is the multitude of different information types (e.g., syntactic, phonological, speaker identity, world knowledge) and sources (e.g., visual, auditory, tactile) that can be rapidly combined at various levels of abstraction in order to arrive at an interpretation of the intended message (e.g., Hagoort & van Berkum, 2007). This implies that whenever we try to make sense of linguistic input, numerous brain regions or systems interact dynamically in order to process and combine these different types of information at multiple hierarchical levels. Investigating neural oscillations provides a window onto exactly this coupling and uncoupling of functional brain networks (e.g., Pfurtscheller & Lopes da Silva, 1999; Singer, 1993; Varela, Lachaux, Rodriguez, & Martinerie, 2001), and it is thus not surprising that this approach is becoming increasingly popular for uncovering various aspects of the cortical dynamics supporting language comprehension (e.g., Friederici & Singer, 2015; Giraud & Poeppel, 2012; Lewis & Bastiaansen, 2015; Strauß, Kotz, Scharinger, & Obleser, 2014; Weiss & Mueller, 2012).

One popular proposal, the 'frequency-based segregation of syntactic and semantic unification' hypothesis (Bastiaansen & Hagoort, 2015), links oscillatory activity in the beta frequency range to syntactic unification operations, and oscillatory activity in the gamma frequency range to semantic unification operations. There is a relatively large body of evidence suggesting a relationship between modulations of beta oscillatory activity and manipulations of syntactic processing (see Lewis, Wang, & Bastiaansen, 2015 for review). A number of studies have shown that beta power was higher for syntactically acceptable sentences compared to sentences containing various types of syntactic violation (Bastiaansen, Magyari, & Hagoort, 2010; Davidson & Indefrey, 2007; Kielar, Meltzer, Moreno, Alain, & Bialystok, 2014; Kielar, Panamsky, Links, & Meltzer, 2015). Another study showed that beta power was higher for long- compared to short-distance subject-verb agreement dependencies at the point in the sentences where the dependency could be resolved (Meyer, Obleser, & Friederici, 2013). When comparing centre-embedded relative clauses to their right-branching counterparts, Bastiaansen and Hagoort (2006) reported higher beta power for the syntactically more complex centre-embedded variety. Finally, higher beta coherence between anterior and posterior electrodes was observed for syntactically more complex object-relative clauses compared to their subject-relative counterparts (Weiss et al., 2005). Importantly, the relative clause sentences in the studies just described were never locally ambiguous (between subject- and object-relative clause sentences), and there was therefore no point within these sentences at which an a-priori preference for a particular sentence structure (e.g., subject-relative) was disconfirmed in favour of a less common sentence structure (e.g., object-relative). These studies together suggest that disrupting syntactic processing leads to a decrease in beta activity, while beta activity is higher when syntactic processing becomes more difficult. Extending these findings, Bastiaansen et al. (2010) showed that the level of beta power increases over the course of a sentence for syntactically acceptable sentences compared to random word lists (see Bastiaansen & Hagoort, 2015 for a replication of this finding), and that for sentences containing a syntactic violation, beta power increases up to the point of the violating word, after which it falls back to baseline levels.

One problem for a strict mapping between beta oscillatory activity and syntactic unification operations is that modulations of beta power have also been observed for manipulations of semantic processing (Kielar et al., 2014, 2015; Luo, Zhang, Feng, & Zhou, 2010; Wang, Jensen et al., 2012), and for disruptions of rhythmical structure (Luo et al., 2010) during language comprehension. Furthermore, the case of Spanish 'Unagreement' shows a decrease in beta power following the target word, even though it does not lead to a syntactic violation (Pérez, Molinaro, Mancini, Barraza, & Carreiras, 2012).

These concerns resulted in the proposal that during language comprehension, just as in other more domain general contexts (Engel & Fries, 2010), oscillatory activity in the beta frequency range might be related to the maintenance or change of the current cognitive set, rather than explicitly to syntactic processing (Lewis & Bastiaansen, 2015; Lewis et al., 2015). Under this proposal, whenever the language comprehension system encounters a cue in the linguistic input indicating that the sentence-level meaning under construction needs to be changed, we should observe a decrease in beta activity in anticipation of the necessary change in the underlying network of regions that interact to represent and construct that sentence-level meaning. Similarly, if the system expects that the current sentence-level meaning needs to be actively maintained, we should observe an increase in beta activity in order to maintain the current network configuration. This proposal (henceforth the ‘beta-maintenance’ hypothesis) can account for all the evidence reviewed above, where for instance syntactic and semantic violations (as well as violations of rhythmical structure and unexpected agreement marking) act as cues to the language comprehension system, indicating that the current sentence-level meaning needs to change, and hence beta power decreases (see Lewis & Bastiaansen, 2015 and Lewis et al., 2015 for discussion). The 'frequency-based segregation of syntactic and semantic unification' hypothesis (henceforth the ‘beta-syntax’ hypothesis) can account for all the beta findings related to syntactic processing that have been discussed, but not those related to disruptions of semantic or rhythmical structure. It may therefore be considered a less general version of the 'beta-maintenance' hypothesis that only applies to beta in relation to syntactic processing.

An important question that has not yet been addressed is: within the domain of syntactic processing, are there cases where the 'beta-maintenance' and the 'beta-syntax' hypotheses make different predictions for modulations of beta activity, and if so, which hypothesis most accurately captures these beta modulations? Additionally, the roles for beta oscillations in language comprehension proposed by the ‘beta-syntax’ hypothesis and the ‘beta-maintenance’ hypothesis have not yet been directly compared with one another. As pointed out by Lewis et al. (2015), the two theories make different predictions about how beta activity should be modulated when the language comprehension system encounters linguistic input that is unexpected (or less expected), yet does not constitute a grammatical violation. In the present study we directly compared the two theories based on this suggestion, employing a syntactic manipulation (locally ambiguous subject- compared to object-relative clause sentences, where disambiguation between the two constructions based on the linguistic input has to wait until the end of the relative clause, but until that point readers have a clear preference for the subject-relative reading) with object-relative clause sentences as the critical test case. This offers a critical comparison between the two hypotheses, and is the first time that the 'beta-maintenance' hypothesis is put directly to the test. Importantly, this comparison is different from previous subject/object relative clause comparisons because in those cases there is no local ambiguity between a subject- or an object-relative clause reading. This means that in those cases, while syntactic complexity (and hence processing demands) is higher for object-relative clause sentences, there is no local disambiguation point within the sentence at which the input disconfirms a preferred syntactic construction and the comprehension system has to change the current sentence-level meaning. As we outlined above, in such cases beta power increases due to the increased syntactic complexity, and consequent need to maintain the current sentence-level meaning under more demanding processing conditions.

Locally ambiguous object-relative clause sentences are interesting in the context of a comparison between the ‘beta-syntax’ and the ‘beta-maintenance’ hypotheses because numerous studies have shown that while they do not contain any grammatical violations, under a number of circumstances they result in processing difficulty compared to subject-relative clause sentences (henceforth the OR-SR processing asymmetry; see Table 5.1 for example sentences). These processing difficulties have been shown in studies using self-paced reading (as well as other related reaction time approaches), eye-tracking, and ERP measures. Additionally, processing difficulties are observed for a variety of languages (although for some languages the processing asymmetry is reversed with subject-relative clause sentences being more difficult to process than object-relative clause sentences), where the difference between subject- and object-relative clause sentences is often realized in different ways (e.g. word order versus inflectional marking): English (e.g., Caplan, Alpert, & Waters, 1998; Caramazza & Zurif, 1976; Ford, 1983; Holmes & O'Regan, 1981; Just, Carpenter, Keller, Eddy, & Thulborn, 1996; King & Just, 1991; Wanner & Maratsos, 1978); Dutch (e.g., Frazier, 1987; Mak, Vonk, & Schriefers, 2002; Mak, Vonk, & Schriefers, 2006); German (e.g., Mecklinger, Schriefers, Steinhauer, & Friederici, 1995; Schriefers, Friederici, & Kuhn, 1995); French (e.g., Cohen & Mehler, 1996; Frauenfelder, Segui, & Mehler, 1980), Basque (e.g., Carreiras, Duñabeitia, Vergara, de la Cruz-Pavía, & Laka, 2010); Chinese (e.g., Gibson & Wu, 2013; Hsiao & Gibson, 2003; Lin, 2008); Hungarian (e.g., MacWhinney & Pléh, 1988); Korean (e.g., Lee, Lee, & Gordon, 2007; Kwon, Gordon, Lee, Kluender, & Polinsky, 2010); Japanese (e.g., Ueno & Garnsey, 2008); Spanish (e.g., Betancort, Carreiras, & Sturt, 2009). There is not yet consensus on exactly which factor(s) result in this OR-SR processing asymmetry, and explanations can be divided into three broad classes: 1) memory/resource-based models; 2) semantic/pragmatic models; 3) frequency-based models (see Gordon & Lowder, 2012 for review).

Memory/resource-based models claim that the OR-SR processing asymmetry comes about because of the extra cognitive burden imposed by maintaining the object referent in memory across more intervening words before it can be integrated at the relative clause verb for the case of object-relative sentences (e.g., Just & Carpenter, 1992; King & Just, 1991; Waters & Caplan, 1996). More recent approaches have emphasized the role of sentence-internal cues that can affect memory processes. The cue-based parsing framework (e.g., Lewis, Vasishth, & Van Dyke, 2006; Lewis & Vasishth, 2005; Van Dyke & Lewis, 2003) for instance suggests that memory interference (caused by similarity between the two noun phrases that could potentially take the role of subject of the relative clause) results in the OR-SR processing asymmetry, after the verb in the relative clause cues retrieval of both potential noun phrases that will take the roles of either subject or object of the relative clause. The dependency locality theory (e.g., Gibson, 1998, 2000; Grodner & Gibson, 2005; Warren & Gibson, 2002) on the other hand claims that object-relative clause sentences are more difficult to process than subject-relative clause sentences because the initial noun phrase from the matrix clause has to be maintained in memory for longer (specifically across more intervening potential discourse referents).

Semantic/Pragmatic models claim that the meaning of a subject-relative clause sentence is derived more straightforwardly than the meaning of an object-relative clause sentence, and that this results in the OR-SR processing asymmetry (e.g., King & Just, 1991). One approach argues that pragmatic and discourse factors are key to the observed asymmetry, suggesting that the function of object-relative clauses is to ground less familiar information in the preceding discourse and to then modify that information with the more familiar noun phrase (e.g., Fox & Thompson, 1990; Gordon & Hendrick, 2004). The OR-SR processing asymmetry arises because the noun phrase within the relative clause does not have a grounding function for subject-relative clause sentences and hence does not require the additional processing within the relative clause that it does in the case of object-relative clause sentences, where it modifies the foregrounded information. Another approach argues that object-relative clause sentences involve a perspective shift as the subject head is modified, and that this perspective shift is not present for subject-relative clause sentences, thus explaining the OR-SR processing asymmetry (e.g., MacWhinney & Pléh, 1988). A related approach explains the asymmetry as a result of the relative animacy of the antecedent noun phrase and the relative clause-internal noun phrase (e.g., Gennari & MacDonald, 2008, 2009; Mak et al., 2002; Mak, Vonk, & Schriefers, 2006; Traxler, Morris, & Seely, 2002; Traxler, Williams, Blozis, & Morris, 2005), where the first animate noun phrase encountered is automatically assigned the role of subject of the relative clause, and this results in processing difficulty for the object-relative clause sentences, where some re-analysis is required to assign the subject role to the alternative noun phrase (Traxler et al., 2002; Traxler et al., 2005).

Frequency-based models appeal to the role of experience with different kinds of sentence structure to explain the OR-SR processing asymmetry (e.g., Reali & Christiansen, 2007; Wells, Christiansen, Race, Acheson, & MacDonald, 2009). The argument is that sentence structures that occur more frequently in a language are easier to comprehend than those that occur only rarely, and that the relative difference in frequency of occurrence between subject- and object-relative clause sentences in most languages can account for the observed OR-SR processing asymmetry.

One approach has concentrated on how the frequency of occurrence of sequences of words from a particular set of syntactic categories within a language (e.g., noun-verb-noun for highly frequent English active sentence) can result in easier comprehension of sequences of words from the same set of syntactic categories in the case of subject-relative clause sentences (noun-verb-noun-verb) compared to sequences of words from a different set of syntactic categories in the case of object-relative clause sentences (noun-noun-verb-verb; e.g., Elman, 1991). Other approaches focus on a reduction in uncertainty about upcoming linguistic input, and how this reduction in uncertainty about upcoming input is higher for subject-relative sentences than for object-relative sentences at the point of disambiguation (e.g., Hale, 2001; Levy, 2008).

Whichever of these models turns out to be correct (perhaps many or all of them for different aspects of relative clause processing), they have in common that at the point of disambiguation within an object-relative clause the language comprehension system encounters an unexpected (or less expected) event. This event could indicate either that some form or re-analysis is required, that more difficult memory retrieval operations will be engaged, that a less frequent sentence construction will be processed, or that the sentence structure implied by the input does not match the predicted structure. We argue that this event constitutes a cue to the language comprehension system indicating that the sentence-level meaning constructed up to that point needs to be changed (i.e., a change in the argument structure). This provides exactly the situation necessary for comparing the predictions of the 'beta-syntax' and 'beta-maintenance' hypotheses, where the linguistic input is unexpected (or less expected) based on the representation of the sentence-level meaning up to that point, but does not constitute a grammatical violation.

Participants read Dutch relative clause sentences like those in Table 5.1 (see also Mak, Vonk, & Schriefers, 2002) while their MEG was recorded. The auxiliary at the end of the relative clause could agree in number with either the antecedent noun phrase in the matrix clause or with the noun phrase within the relative clause, resulting in either a subject-relative (SR condition) or an object-relative (OR condition) clause reading of the sentence. Both the antecedent noun phrase and the relative clause-internal noun phrase had animate referents. The auxiliary could also fail to agree in number with either noun phrase, resulting in a grammatical violation (AVR condition) at the end of the relative clause. Importantly, Dutch readers show a clear preference for a subject-relative reading of such sentences. The object-relative clause sentences occur less frequently according to a corpus analysis, and they result in processing difficulties at the disambiguating auxiliary within the relative clause (Mak et al., 2002). This means that both the OR and AVR sentences are relatively less expected for a typical Dutch reader, while only the AVR condition constitutes a grammatical violation. We performed a time-frequency analysis of power changes relative to a baseline period immediately preceding the target word in a frequency range from 2 to 30 Hz. This allows us to test whether beta power at the target word is modulated differently or in a comparable manner for the OR and AVR conditions, compared with the SR condition where the target word is not unexpected.

Table 5.1 Example materials for the main conditions and their English translation (in italics).

Condition	*Example Materials*
SR	Achteraf praat de vader, die de zonen bij het concert bewonderd heeft, met de dirigent over het optreden.
	Afterwards discusses the father, that the sons at the concert admired has, with the conductor about the performance.
OR	Achteraf praat de vader, die de zonen bij het concert bewonderd hebben, met de dirigent over het optreden.
	Afterwards discusses the father, that the sons at the concert admired have, with the conductor about the performance.
AVR	* Achteraf praat de vader, die de zonen bij het concert bewonderd hebt, met de dirigent over het optreden.
	** Afterwards discusses the father, that the sons at the concert admired have, with the conductor about the performance.*

Notes: SR: subject-relative clause condition; OR: object-relative clause condition; AVR: agreement violation within relative clause condition; auxiliary verb and referent that agrees with it in both grammatical person and number are underlined

We hypothesized that beta power at the target word would be modulated for both the OR and the AVR conditions compared to the SR condition. If the 'beta-syntax' hypothesis is correct, then beta power should be higher for the OR condition than for the SR condition, while it should be lower for the AVR condition than for the SR condition. If on the other hand the 'beta-maintenance' hypothesis is correct, beta power should be lower for both the OR and the AVR conditions compared to the SR condition. Put differently, according to the 'beta-maintenance' hypothesis beta power for the OR and AVR conditions should pattern together, while according to the 'beta-syntax' hypothesis beta power should increase for the OR condition (increased syntactic complexity) and decrease for the AVR condition (syntactic violation).

In addition to the time-frequency analysis of power, we also performed an event-related field (ERF) analysis in order to ensure that our stimuli give rise to the typically observed magnetic P600 effect for syntactic violations (e.g., Bastiaansen et al., 2010; Service, Helenius, Maury, & Salmelin, 2007), and to investigate whether a similar magnetic P600 effect is observed when comparing OR and SR sentences.

5.2 Methods

5.2.1 Participants

Thirty native speakers of Dutch took part in the experiment, 24 of whom were included in the final analysis (3 males, 21 females; aged 18 to 35). Participants provided informed consent and were paid or equivalently rewarded with course credits for their participation. All participants reported normal or corrected-to-normal vision, and were right handed. None of the participants reported any neurological impairment. Three participants were excluded from the final analysis due to poor performance on the comprehension questions (less than 65 % correct answers overall). One further participant was excluded due to recording problems, and another 2 participants were excluded in order to balance the number of participants who were assigned to each experimental list (for lists with too many participants, those participants with the worst performance on OR comprehension questions were excluded).

5.2.2 Stimulus materials

All stimuli consisted of Dutch relative clause sentences, each between 11 and 22 words long. For the main experimental materials (Table 5.1), the relative clause always consisted of the relative pronoun *die* (English *that*), followed by a full noun phrase (NP), then by a prepositional phrase, then by a past participle, and finally by an auxiliary verb. The relative clause was always preceded by an antecedent NP together with some modifier, and followed by at least 3 words to complete the matrix clause of the sentence. Conditions differed in terms of whether the auxiliary at the end of the relative clause (the target word for the main experimental conditions) agreed in grammatical number with the antecedent NP (SR condition), with the NP within the relative clause (OR condition), or did not agree in grammatical person with either NP, thus resulting in a grammatical violation within the relative clause (AVR condition). The referents of both the antecedent and the relative clause-internal NPs were animate, and the past participle in the relative clause was not biased in terms of which NP was more likely to be the grammatical subject of the relative clause. Up to the point of the auxiliary (the target word for the main materials) in the relative clause, these sentences are ambiguous in terms of whether they should be read as a subject-relative clause, as an object relative clause, or whether they constitute a grammatical agreement violation.

Table 5.2 Example materials for the filler conditions and their English translations (in italics).

Condition	*Example Materials*
AGR	De vrolijke clown, die heel hard lacht, werpt de hoed naar het meisje.
	The merry clown, that very loudly laughs, throws the hat to the girl.
AV	* De vrolijke clown, die heel hard lacht, werpen de hoed naar het meisje.
	* *The merry clown, that very loudly laughs, throw the hat to the girl.*

Notes: AGR: subject-verb agreement condition; AV: agreement violation outside relative clause condition; verb and referent that agrees with it in grammatical number are underlined

Two additional conditions where the relative clause was unambiguously subject-relative (no NP was present within the relative clause) were included as fillers (Table 5.2). The relative clause was always preceded by an antecedent NP, and followed by at least 3 words to complete the main clause of the sentence. For the fillers, the matrix clause verb directly following the relative clause (the target word for the filler conditions) was inflected to either agree (AGR condition) or not agree (resulting in a grammatical violation; AV condition) in grammatical number with the subject of the sentence. These filler conditions provide a contrast between grammatically acceptable sentences and sentences containing a grammatical violation in the relatively less complex context of unambiguously subject-relative clause sentences. This contrast was used to select regions of interest for the statistical comparison of the main experimental conditions (see *Selection of regions of interest* below).

For the SR condition 120 sentences were constructed according to the specifications just described. About a quarter of the sentences were taken directly from a self-paced reading and eye-tracking study by Mak, Vonk, and Schriefers (2008), while the remainder were adapted from subject- and object-relative clause sentences used in an unpublished study. For half of the sentences the antecedent NP was singular while the NP within the relative clause was plural, and vice versa for the other half. One hundred and twenty sentences for the OR condition were constructed by switching the auxiliary in the relative clause from the SR sentences (i.e., *heeft* became *hebben* and *hebben* became *heeft*) so that it agreed in grammatical number with the relative clause-internal NP rather than with the antecedent NP. To create grammatical person agreement violations in the 120 sentences for the AVR condition, the auxiliary within the relative clause (*heeft* or *hebben*) was replaced by the Dutch auxiliary *hebt*, which carries second person singular grammatical marking and therefore does not agree in person with either the antecedent NP or the relative clause-internal NP.

For the AGR filler condition 80 sentences were constructed according to the specifications described above for the filler sentences. The antecedent NP for half the sentences was singular (and thus in order for the sentence to be grammatical so was the inflectional marking on the verb in the matrix clause) and for the other half it was plural (again with plural inflectional marking on the matrix clause verb). To create grammatical number agreement violations in the 80 sentences for the AV filler condition, singular matrix clause verbs from the AGR condition were replaced by verbs with plural inflectional marking and plural matrix clause verbs were replaced by verbs with singular inflectional marking.

Participants saw 40 sentences from each of the conditions over the course of the experiment. Which of the 120 sentences from each of the main conditions and which of the 80 sentences from each of the filler conditions were presented was separately counterbalanced across participants, such that participants never saw the same sentence more than once throughout the experiment. Half of the sentences presented from each condition had a singular antecedent NP and plural relative clause-internal NP, and vice versa for the other half.
Resulting experimental lists were then pseudo randomized according to the following criteria: 1) no more than two consecutive presentations of a sentence from the same experimental condition; 2) repetition of sequences of 5 or more sentences from any particular sequence of conditions was avoided.

5.2.3 Experimental design and procedure

Participants were tested in a dimly lit, sound-attenuating, magnetically and electrically shielded room. They were seated in front of a display, with a viewing distance of approximately 90 cm. The display consisted of a back-projection screen inside the magnetically shielded room, on which all stimuli were presented using a set of mirrors and an LCD projector positioned outside the magnetically shielded room in order to minimize electrical interference. Letters were presented in black on a dark grey background using a 20-point sized Consolas font type.

Sentences were presented word by word in the centre of the screen. For each sentence, the first letter of the first word was capitalized, the word directly preceding the relative clause and the last word of the relative clause were presented followed by a comma, and the final word of the sentence was presented with a period. A single trial consisted of a sentence, a movement cue, and a fixation cross (and sometimes a comprehension question). Words were presented for between 300 and 400 ms (randomly chosen for each word), followed by a blank screen between words presented for between 100 and 200 ms. The stimulus onset asynchrony (SOA) between two words was always 500 ms (e.g., if the word was presented for 325 ms then the blank screen would last for 175 ms)[1]. Each trial began with the presentation in the centre of the screen of three asterisks two spaces apart for 3000 ms, indicating that participants could move their eyes and blink. This was immediately followed by a fixation cross presented in the centre of the screen for 1500 ms, indicating that eye movements and blinking should be avoided, and that the sentence was about to start. The first word of the sentence immediately followed the fixation cross. Each sentence lasted between 5500 and 11000 ms and a single trial lasted between 10000 and 15500 ms.

Participants were instructed to read all sentences attentively for comprehension, and that every once in a while they might notice a grammatical error, but should continue reading to the end anyway. They read a total of 200 sentences (40 SR, 40 OR, 40 AVR, 40 AGR, and 40 AV), presented in 20 blocks of 10 sentences each, with self-timed breaks between blocks. After a random 10 % of the sentences (4 from each of the conditions) a comprehension question appeared on the screen instead of the next trial. Participants were required to provide a response with the index finger ('yes' response) or middle finger ('no' response) of their right hand, indicating whether the statement on the screen correctly described the sentence they had just read. The question remained on the screen for 6500 ms or until participants made a response, after which the next trial began. Whether or not a statement correctly described the sentence just read was counterbalanced across participants (2 'yes' and 2 'no' responses to the 4 questions from each condition). Ten training sentences (not used in the main experiment) were presented to participants before the experiment began.

5.2.4 MEG recordings

Participants were seated upright in the MEG system with their heads as close as possible to the inside of the helmet. MEG signals were recorded from a whole-head MEG system with 275 axial gradiometers (CTF MEG systems, VSM MedTech) at a sampling rate of 1200 Hz and with a 300 Hz low-pass anti-aliasing filter. Participants' head position relative to the helmet was monitored in real-time (Stolk, Todorovic, Schoffelen, & Oostenveld, 2013) using 3 localization coils, one placed on participants' nasion and one in each ear canal. After each block participants were asked to reposition their head in case of a deviation from their original head position exceeding 10 mm. Bipolar electrode montages were used to record participants' electrocardiogram, as well as their horizontal (electrodes positioned at outer canthi) and vertical (electrodes positioned above and below left eye) electrooculograms. Electrode impedance was kept below 20 kΩ.

5.2.5 Data pre-processing

MEG data were analysed using the FieldTrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011) running in a MatLab environment (R2014b; Mathworks, Inc.). MEG and EEG channels were high-pass filtered above 0.1 Hz, and a band-stop filter was applied at 50, 100, and 150 Hz (all using a windowed sinc finite-impulse response filter with FieldTrip default settings) in order to minimize the effects of power line noise (50 Hz). Data were then segmented from 2000 ms prior to the onset of the first word to 1500 ms after the onset of the last word of each sentence for all conditions together. In three separate rounds of artifact rejection, the data were temporarily transformed (filtered and/or normalized) to ensure that it was highly likely that different types of well known artifacts could be detected. Segments containing superconducting quantum interference device (SQUID) jump artifacts, muscle artifacts, and eye blink artifacts were then removed by visual inspection and partial rejection (only the artifactual portion of a sentence segment was removed, after accounting for edge effects).

Next, data from the MEG channels were temporarily downsampled to 300 Hz and decomposed into independent components (ICA using the 'runica' implementation in FieldTrip with default settings), requesting the 50 component time-courses accounting for the highest variance in the data. Components which captured residual eye-blinks, eye movements (including obvious microsaccadic components; Hipp & Siegel, 2013), or cardiac response were removed from the data (Jung et al., 2000; Makeig, Jung, Bell, Ghahremani, & Sejnowski, 1997). Between 2 and

10 components were removed per participant. Data were then segmented from -1000 to 1500 ms relative to target word onset for all conditions together.

5.2.6 Event-related fields analysis

Single-trial pre-processed data for each participant were low-pass filtered below 30 Hz (using a windowed sinc finite-impulse response filter with FieldTrip default settings), and a baseline correction was performed using a period from -200 to 0 ms relative to target word onset. Next, data were segmented into trials from the SR (*M* = 33.64, SD = 5.03), OR (*M* = 33.96, SD = 4.85), and AVR (*M* = 33.64, SD = 4.72) conditions from -200 to 1500 ms relative to target word onset, and averaged over trials within each condition to obtain participant-specific axial gradiometer representations of the ERF waveforms for each condition. Planar gradient representations of the data from each condition were then estimated for each MEG sensor based on neighbouring sensors (within a distance of 6 cm or less from the sensor of interest). First the spatial derivatives of each sensor in two orthogonal directions were computed, and then their absolute values were averaged using the Pythagorean theorem (see Bastiaansen & Knösche, 2000 for more details about the procedure). Finally, a condition-specific baseline correction was applied for each condition from -200 to 0 ms relative to target word onset.

5.2.7 Time-frequency analysis

Single-trial pre-processed data for each participant were DC-offset corrected using the entire segment from -1000 to 1500 ms relative to target word onset. Next the spatial derivatives of each sensor in two orthogonal directions were computed using neighbouring sensors (within a distance of 6 cm or less from the sensor of interest). TF analyses were carried out on this representation of the data so that in a later step the planar gradient representation of the TF data could be estimated (Bastiaansen & Knösche, 2000). Data were then segmented into trials from the SR (*M* = 33.64, SD = 5.03), OR (*M* = 33.96, SD = 4.85), AVR (*M* = 33.64, SD = 4.72), AGR (*M* = 33.96, SD = 3.87), and AV (*M* = 33.68, SD = 5.21) conditions from -1000 to 1500 ms relative to target word onset.

A multitaper approach (Mitra & Pesaran, 1999) was used to compute TF representations of power for the single trial data of each participant. Since we were interested in differences in beta power we chose a TF tradeoff that would result in our analysis being highly sensitive to fluctuations of power in the beta frequency range. Time-resolved power spectra of the data between 2 and 30 Hz were computed using a sliding window approach. Sliding windows of 500 ms were applied in frequency steps of 2 Hz and time steps of 20 ms across the entire time axis. Frequency smoothing at 4 Hz was achieved using a sequence of Slepian tapers (Mitra & Pesaran, 1999), and the power estimate at each time point is thus based on data points from the preceding and following 250 ms. Single-trial power spectra were then averaged within each condition. Finally, to obtain the planar gradient representation of the TF data from each condition the absolute values of the power spectra for the two spatial derivatives at each sensor were summed (Bastiaansen & Knösche, 2000). This resulted in a condition-specific TF representation of power for each participant. For each condition these participant averages were then expressed as a relative change (in dB) from a baseline period between 600 ms and 100 ms prior to the onset of the target word.

5.2.8 Selection of regions of interest

The filler conditions in this experiment (AGR and AV) allowed us to contrast grammatical relative clause sentences with relative clause sentences containing a syntactic violation. Different from our main experimental manipulations, the syntactic violation occurs outside the relative clause (in the matrix clause) and in an unambiguous sentence context. It is also a violation of grammatical number agreement, rather than of grammatical person agreement as is the case for the AVR condition. We used this contrast (AGR - AV) to select regions of interest in the beta frequency range (Figure 5.1) for testing our main hypothesis. First, dependent samples T-values were computed to compare the AGR and AV conditions for every sensor-time-frequency triplet in the data, and the resultant T-values were thresholded at positive or negative 2. Next, we plotted the topographical representation of these thresholded T-values (Figure 5.1A), averaged over the entire time interval (0 to 1200 ms relative to target word onset) and over the beta frequency range (16 to 26 Hz; the beta frequency range is typically taken from 12 to 30 Hz but our selection was made taking into account the 4 Hz frequency smoothing that was used). The figure clearly shows one cluster of sensors exhibiting high T-values over the left hemisphere (MLF14, MLF25, MLF35, MLF46, MLF56, MLT11, MLT12, MLT21; marked in the figure), and another cluster over the right hemisphere (MRC14, MRC15, MRC16, MRF54, MRF55, MRF63, MRF64, MRF65, MRT22, MRT32, MRT41; marked in the figure). We then separately plotted TF representations

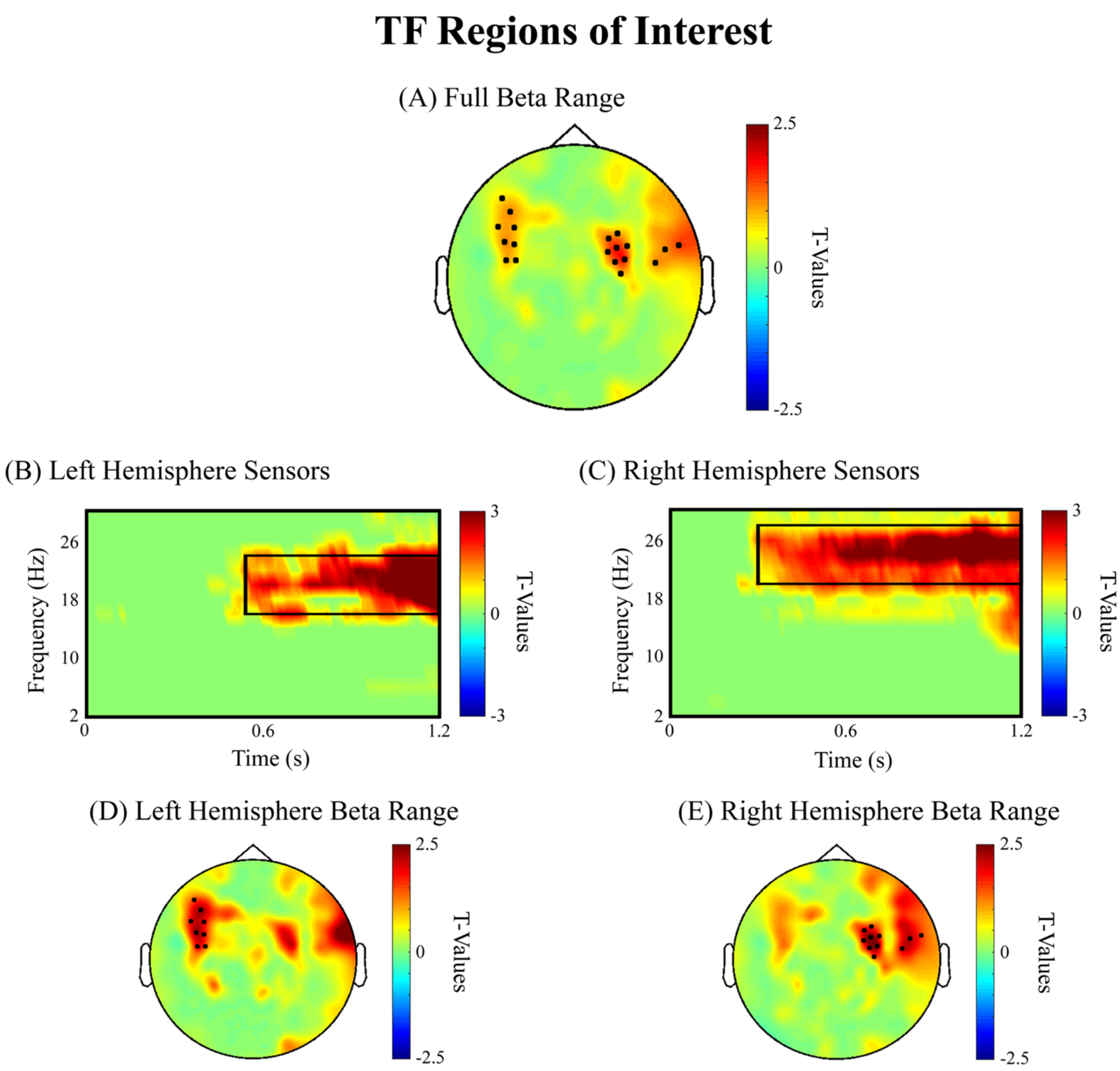

Figure 5.1 Selection of regions of interest. (A) Topographical representation of T-values for the comparison between the filler conditions (AGR - AV) thresholded at +/-2 and averaged over the beta frequency range (16 to 24 Hz) and over the entire time interval (0 to 1200 ms relative to target word onset). (B) TF representation of thresholded T-values averaged over left hemisphere sensors of interest (MLF14, MLF25, MLF35, MLF46, MLF56, MLT11, MLT12, MLT21). (C) TF representation of thresholded T-values averaged over right hemisphere sensors of interest (MRC14, MRC15, MRC16, MRF54, MRF55, MRF63, MRF64, MRF65, MRT22, MRT32, MRT41). (D) Topographical representation of thresholded T-values averaged over the refined time (540 to 1200 ms relative to target word onset) and frequency (16 to 24 Hz) ranges for the left hemisphere region of interest. (E) Topographical representation of thresholded T-values averaged over the refined time (300 to 1200 ms relative to target word onset) and frequency (20 to 28 Hz) ranges for the right hemisphere region of interest. Sensors of interest marked with black dots in (A), (D), and (E); TF regions of interest indicated with black boxes in (B) and (C).

of the T-values averaged over those two clusters of sensors (Figures 5.1B and 5.1C). This revealed different (but overlapping) time and frequency ranges for the left (540 to 1200 ms relative to target word onset; 16 to 24 Hz) and right (300 to 1200 ms relative to target word onset; 20 to 28 Hz) hemisphere sensor clusters (black boxes in Figures 5.1B and 5.1C). Plotting the topographical representation of the T-values for these refined TF ranges (Figures 5.1D and 5.1E) confirmed that differences were more pronounced for the restricted ranges, and we therefore selected these two regions of sensors, time, and frequency points for statistical comparison of the main experimental conditions (SR, OR, and AVR).

5.2.9 Statistical analyses

The statistical significance of all ERF comparisons was evaluated using a cluster-based random permutation approach (Maris & Oostenveld, 2007). We used this approach because of its natural handling of the multiple comparisons problem. Cluster-based random permutation statistics control the family-wise error rate by making use of the spatial, and temporal autocorrelation in MEG data. In short, a dependent-samples T-test is performed for every data point (sensor-time point) giving uncorrected p-values. A pre-set significance level is chosen (here 5% two-tailed) and any data points not exceeding this level are discarded (set to zero). Clusters are calculated from the remaining non-zero data points based on their adjacency in space (adjacent sensors) and time.

Cluster-level statistics are then calculated by summing the values of the T-statistics for all data points in each cluster. A permutation distribution is created by randomly assigning participant averages to one of the two conditions 3000 times, and each time calculating cluster-level statistics as just described. The highest cluster-level statistic from each randomization is entered into the permutation distribution and the cluster-level statistics calculated for the measured data are compared against this distribution. Clusters falling in the highest or lowest 2.5th percentile of the estimated distribution were considered significant (although P-values reported are corrected for the 2 tests performed and are considered significant at $p < 0.05$). We separately compared the OR with the SR, the AVR with the SR, and the AVR with the OR conditions for the entire time interval of 0 to 1200 ms relative to target word onset.

For the TF data we subjected the mean power from both the left (540 to 1200 ms; 16 to 24 Hz; sensors MLF14, MLF25, MLF35, MLF46, MLF56, MLT11, MLT12, MLT21) and right (300 to 1200 ms; 20 to 28 Hz; sensors MRC14, MRC15, MRC16, MRF54, MRF55, MRF63, MRF64, MRF65, MRT22, MRT32, MRT41) hemisphere regions of interest to a repeated measures ANOVA, with condition (SR, OR, AVR) as a fixed within group factor. Significant main effects were broken down using simple contrasts, with the SR condition used as control. According to Machley's test sphericity was not violated so no correction of reported values was necessary. A cluster-based random permutation approach was not necessary here, as regions of interest had already been identified based on the AGR-AV contrast. Furthermore, because separate left- and right-hemisphere regions of interest were identified (with different time and frequency ranges in addition to different contributing sensors) separate ANOVAs were performed for each region of interest.

5.3 Results

Participants were excluded from further analysis when they answered less than 65 % of comprehension questions correctly overall. Those participants included in the final analyses scored on average 77 % correct for the comprehension questions (SD = 10 %; Range 65 to 95 %). This suggests that participants were paying attention to the stimuli and understood the sentences they were reading.

5.3.1 Event-related field results

ERF effects were quantified by temporally and spatially contiguous time points identified by the cluster-based permutation approach. Figure 5.2 shows the planar gradient ERF waveforms (left) at a representative sensor for the OR vs SR (Figure 5.2A; sensor MLF35), the AVR vs SR (Figure 5.2B; sensor MLF35), and the AVR vs OR (Figure 5.2C; MRP43) comparisons, along with the topographical representation of the difference between conditions averaged over a time interval selected based on the statistical output (right).

We observed a statistically significant positive difference between the OR and SR conditions ($p = 0.02$) over 82 sensors, with the time region contributing to the effect between 493 and 705 ms (shaded region of the waveform in Figure 5.2A) relative to target word onset. Between about 480 and 850 ms relative to target word onset the waveform for the OR condition is more positive than that for the SR condition. The effect is most pronounced over left frontal and left parietal sensors, but is also present over left temporal sensors. Based on the timing and the fact

ERF Results

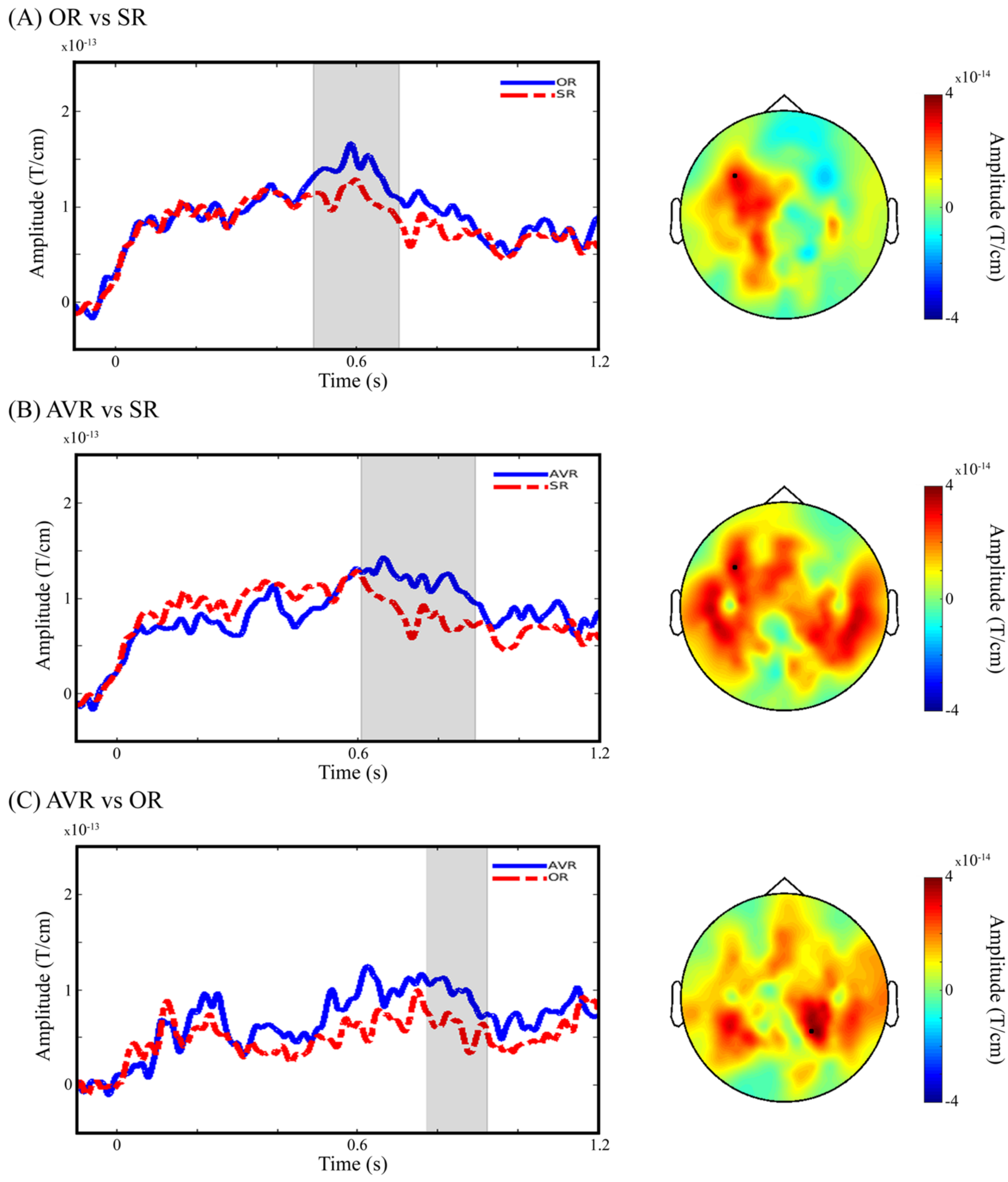

Figure 5.2. Results of the ERF analysis. (A) ERF waveforms for the OR and SR conditions at a representative left hemisphere sensor MLF35 (left) and topographical representation of the difference between the two conditions (OR minus SR) averaged over the time interval contributing to the statistically significant difference (right). (B) ERF waveforms for the AVR and SR conditions at a representative left

hemisphere sensor MLF35 (left) and topographical representation of the difference between the two conditions (AVR minus SR) averaged over the time interval contributing to the statistically significant difference (right). (C) ERF waveforms for the AVR and OR conditions at a representative right hemisphere sensor MRP43 (left) and topographical representation of the difference between the two conditions (AVR minus OR) averaged over the time interval contributing to the statistically significant difference (right). Time interval contributing to the statistically significant difference shaded in grey in waveform plots; representative sensor marked with black dot in topographical representation plots.

that the effect is present over left frontal and temporal sensors, we identify this as the magnetic equivalent of a P600 effect.

For the comparison between the AVR and SR conditions there was a statistically significant positive difference ($p = 0.01$) over 197 sensors, with the time region contributing to the effect between 608 and 892 ms (shaded region of the waveform in Figure 5.2B) relative to target word onset. The waveform for the AVR condition is more positive than that for the SR condition between about 600 and 900 ms relative to target word onset, and the effect is strongest over bilateral temporal and left frontal sensors. We again identify this as a P600 effect, as right temporal sensors have also exhibited such effects in previous MEG studies investigating grammatical violations (e.g., Service, Helenius, Maury, & Salmelin, 2007).

Finally, we observed a statistically significant positive difference between the AVR and OR conditions ($p = 0.03$) over 120 sensors, with the time region contributing to the effect between 774 and 925 ms (shaded region in the waveform in Figure 5.2C) relative to target word onset. The waveform for the AVR condition is more positive than that for the OR condition at various time points after about 300 ms relative to target word onset, but the region between about 770 and 920 ms appears to be the most consistent. The effect is strongest over right parietal sensors, although it is also present over left parietal and right temporal sensors. Although the timing is appropriate, based on the topographical representation of the difference we are hesitant to label this a classical P600 effect (see *Magnetic P600 effects* below).

5.3.2 Time-frequency results

We selected two regions of interest in the beta frequency range for statistical comparison, a left (540 to 1200 ms; 16 to 24 Hz; sensors MLF14, MLF25, MLF35, MLF46, MLF56, MLT11, MLT12, MLT21) and a right (300 to 1200 ms; 20 to 28 Hz; sensors MRC14, MRC15, MRC16, MRF54, MRF55, MRF63, MRF64, MRF65, MRT22, MRT32, MRT41) hemisphere region. There were no statistically significant effects for the right hemisphere region of interest. For the left hemisphere region of interest Figure 5.3A shows the TF representation of power at a representative sensor (MLF46) for the OR, SR, and AVR conditions (left), as well as for the difference between the SR and OR and between the SR and AVR conditions (right). The TF range tested is marked by black boxes. Figures 5.3B and 5.3C show the topographical representation of the differences in power between the SR and OR, and between the SR and AVR conditions respectively, averaged over the left hemisphere TF range of interest (the representative sensor used for plotting in Figure 5.3A is marked).

For this left hemisphere region of interest, the repeated measures ANOVA revealed a significant main effect of condition, $F(2,46) = 3.764$, $MSE = 0.112$, $p = 0.031$. Follow-up simple contrasts with the SR condition as control revealed that the SR condition was significantly different from both the OR condition (Figure 5.3A right top and Figure 5.3B), $F(1,23) = 8.271$, $MSE = 0.159$, $p = 0.009$, and the AVR condition (Figure 5.3A right bottom and Figure 5.3C), $F(1,23) = 5.161$, $MSE = 0.234$, $p = 0.033$. These effects are driven by a decrease in beta power within the TF range of interest for both the OR and the AVR conditions, with no change in beta power in this range for the SR condition (Figure 5.3A). The topographical representations of the power differences suggest that both effects are strongly left lateralized, with statistical significance confirmed only for left frontal and left temporal sensors from the left hemisphere region of interest.

5.4 Discussion

The 'beta-syntax' hypothesis (Bastiaansen & Hagoort, 2015) links oscillatory activity in the beta frequency range to syntactic unification operations. On the other hand, the 'beta-maintenance' hypothesis (Lewis & Bastiaansen, 2015; Lewis, Wang, & Bastiaansen, 2015) argues that the experimental evidence linking beta to sentence-level language comprehension is better described under the more domain-general proposal that oscillatory activity in the beta frequency range is related to the maintenance or change of the current cognitive set. We directly compared these two hypotheses by investigating how MEG beta power is modulated when linguistic information is encountered that is unexpected (or less expected) but does not constitute a grammatical violation. Participants read locally ambiguous Dutch relative clause sentences that were disambiguated at the target word as either a subject-relative clause (the preferred grammatical structure for Dutch readers), an object-relative clause (the less common and unexpected structure for Dutch readers), or as containing a syntactic violation.

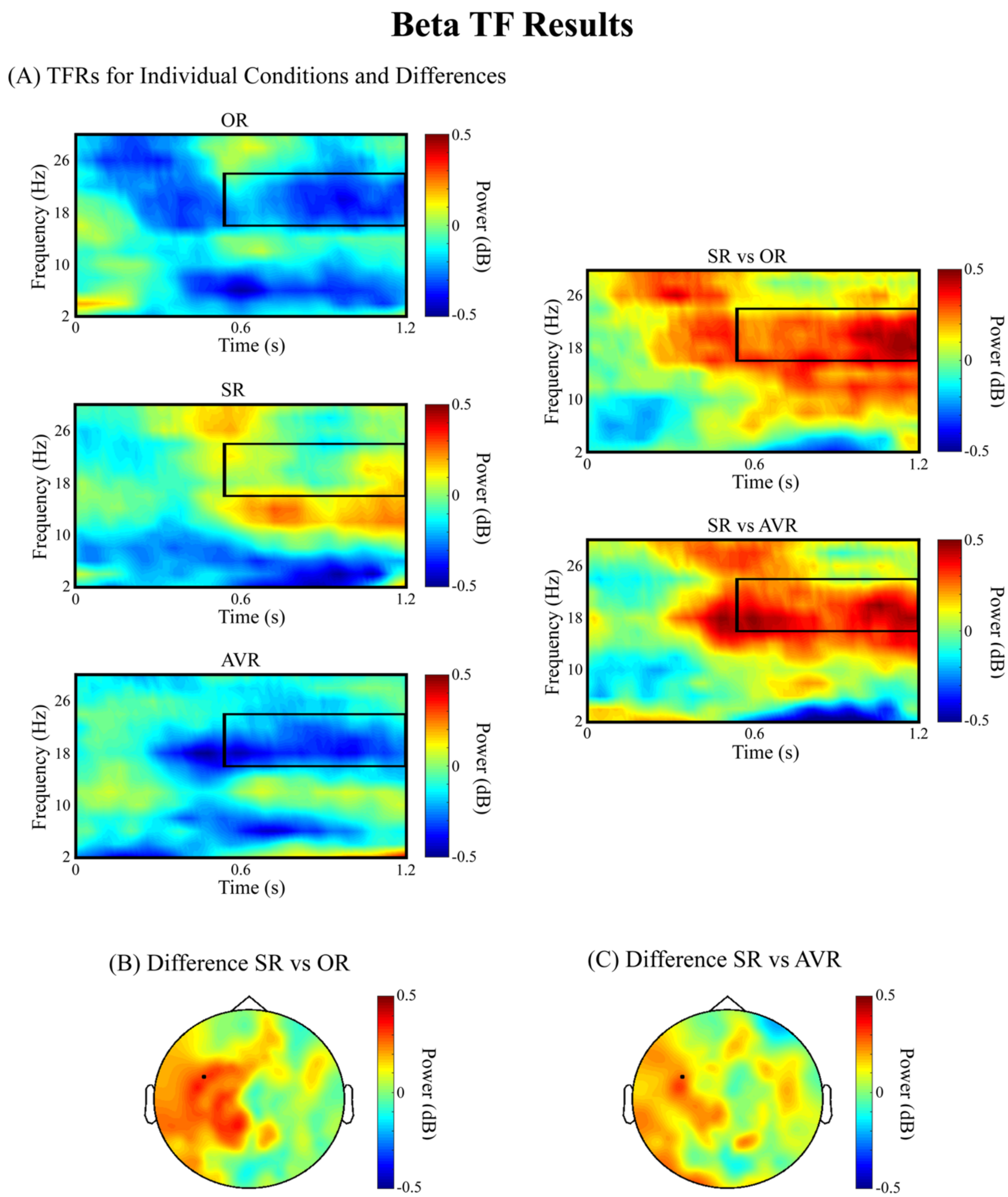

Figure 5.3. Results of TF analysis for left hemisphere region of interest. (A) TF representations of power for the OR (left top), SR (left middle), and AVR (left bottom) conditions, as well as for the differences between the OR and SR (right top) and between the AVR and SR (right bottom) conditions at a representative sensor MLF46. Black boxes indicate the TF range for the left hemisphere region of interest. (B) Topographical representation of the difference in power between the SR and OR conditions averaged over the time (540 to 1200 ms relative to target word onset) and frequency (16 to 24 Hz) ranges of interest. (C) Topographical representation of the difference in power between the SR and AVR conditions averaged over the time (540 to 1200 ms relative to target word onset) and frequency (16 to 24 Hz) ranges of interest. Representative sensor MLF46 marked with black dot in (B) and (C).

The ERF analysis produced a P600 effect when comparing the OR to the SR condition and when comparing the AVR to the SR condition, as well as a late positive difference when comparing the AVR to the OR condition (Figure 5.2). The time-frequency analysis of power produced a significant beta difference for our left hemisphere region of interest when comparing the SR with the OR and the SR with the AVR conditions. Beta power (16 to 24 Hz) over left frontal and left temporal sensors decreased for both the OR and AVR conditions, but was unchanged for the SR condition, in a time window between 540 and 1200 ms relative to target word onset (Figure 5.3).

5.4.1 Magnetic P600 effects

Our P600 effect for the comparison between the AVR and SR conditions confirms previous ERP (e.g., Friederici & Meyer, 2004; Gunter, Stowe, & Mulder, 1997; Hagoort, 2003; Neville, Nicol, Barss, Forster, & Garrett, 1991; Osterhout, Holcomb, & Swinney, 1994; Osterhout & Holcomb, 1992; Rossi, Gugler, Hahne, & Friederici, 2005) and ERF (e.g., Bastiaansen, Magyari, & Hagoort, 2010; Service, Helenius, Maury, & Salmelin, 2007) results, exhibiting a more positive waveform when the target word resulted in a grammatical violation than when it did not. The effect was present between about 600 and 900 ms relative to the onset of the target word. Furthermore, the presence of an effect at bilateral sensors is in good agreement with previously reported magnetic P600 results (Service et al., 2007), where equivalent current dipoles (ECDs) in both the left and right posterior temporal cortex exhibited P600 effects.

Our P600 effect for the comparison between the OR and the SR conditions confirmed previous ERP results (e.g., Friederici, 2002; Holle et al., 2012; Kaan, Harris, Gibson, & Holcomb, 2000), but as far as we are aware is the first report of a magnetic P600 effect associated with an OR-SR processing asymmetry. The OR sentences exhibited a more positive waveform at the target word than the SR sentences between about 500 and 700 ms relative to target word onset. Unlike for the syntactic agreement violation sentences, the topographical representation of the P600 effect for the comparison between OR and SR sentences indicates that the effect was only present at left hemisphere sensors. This suggests that the cortical generators for the P600 response to syntactic violations may be different from those for the P600 response to OR-SR processing asymmetries (although there appears to be considerable overlap), with right hemisphere generators engaged only for syntactic violations. The results we report are however based on sensor level data and so do not allow us to make any strong claims about cortical sources. Future studies should use source localization techniques to follow up on this potential dissociation. That the onset of the P600 effect for the comparison between the OR and SR conditions occurs about 50 ms earlier than that for the comparison between the AVR and SR conditions is also in good agreement with previous ERP findings (see Gouvea, Phillips, Kazanina, & Poeppel, 2010 for review and discussion).

We also observed a statistically significant difference between the AVR and OR conditions that was most prominent over right hemisphere parietal sensors. We suggest that this difference is likely due to the additional recruitment of right hemisphere cortical sources for the AVR condition compared to the OR condition. A question then arises about exactly what the role of these additional right hemisphere cortical sources is. One possibility is that this reflects the recruitment of additional attentional resources for the AVR condition once the language comprehension system 'realizes' that a grammatical parse of the sentence will not be possible. For the OR condition the language comprehension system is able to successfully recover from the unexpected disambiguating auxiliary and make sense of the sentence, so fewer additional attentional resources are necessary. On the other hand, for the AVR condition a successful parsing of the sentence is not possible and so the system may recruit additional attentional resources as a final attempt to recover meaning. This would also explain why the P600 effect is present at both left and right hemisphere sensors when comparing the AVR and SR conditions, but only at left hemisphere sensors when comparing the OR and SR conditions. Our explanation is of course highly speculative and warrants further investigation, but on the other hand right parietal cortex does form part of the dorsal attention network (Petersen & Posner, 2012), and our explanation is in line with accounts of the P600 as a delayed or late form of P3b (e.g., Coulson, King, & Kutas, 1998; Sassenhagen, Schlesewsky, & Bornkessel-Schlesewsky, 2014).

Many previous studies have observed a left anterior negativity (LAN) or early left anterior negativity (ELAN) effect for syntactic violations compared to control sentences (Angrilli et al., 2002; Barber & Carreiras, 2005; Friederici, Pfeifer, & Hahne, 1993; Gunter, Friederici, & Schriefers, 2000; King & Kutas, 1995; Neville et al., 1991). Additionally, an anterior negativity has been observed using MEG for morphosyntactic violations (Service et al., 2007). We might therefore have expected to observe a LAN or ELAN effect in our comparison between the AVR and SR conditions, but we made no such observation. There are however a number of studies comparing syntactic violations with control sentences that do not find LAN or ELAN effects (Gunter et al., 1997; Hagoort, Brown, & Groothusen, 1993; Lau, Stroud, Plesch, & Phillips, 2006;

Nevins, Dillon, Malhotra, & Phillips, 2007; Vos, Gunter, Kolk, & Mulder, 2001). Since these effects do not appear to be consistently observed when comparing syntactic violation sentences with controls, and since our main focus with the ERF analysis was the P600, we do not discuss the absence of LAN or ELAN effects further.

5.4.2 Lower beta power for syntactic violation and object-relative clause sentences

The results of our region of interest analysis provide strong support for the 'beta-maintenance' hypothesis, and against the 'beta-syntax' hypothesis. At left frontal and left temporal sensors between 540 and 1200 ms relative to the onset of the target word, beta power (16 to 24 Hz) decreased for both the AVR and the OR conditions, while in the same region and time range beta power was unchanged from baseline for the SR condition. Both hypotheses predicted that beta power should be lower for syntactic violations (AVR condition) compared to grammatically acceptable control sentences (SR condition). However, the 'beta-maintenance' hypothesis correctly predicted a beta power decrease for the syntactically acceptable but less preferred or unexpected sentences (OR condition), while the 'beta-syntax' hypothesis predicted instead that beta power should increase with the increasing syntactic unification demands related to resolving the processing difficulty.

Two alternative possibilities for the role of oscillatory activity in the beta frequency range during language processing have been proposed (Weiss & Mueller, 2012). First, beta activity has been implicated in the processing of action semantic content during language comprehension (e.g., Moreno et al., 2015; Moreno, de Vega, & León, 2013; van Elk, van Schie, Zwaan, & Bekkering, 2010). We can rule this out immediately as a potential explanation for our findings, since we did not systematically manipulate the action content of our sentences. Second, beta activity has been linked to the maintenance and manipulation of information in working memory (e.g., Onton, Delorme, & Makeig, 2005; Tallon-Baudry, Kreiter, & Bertrand, 1999). Support for such a link comes from biophysically realistic computational modelling (Kopell, Whittington, & Kramer, 2011), which suggests that information can be held online for extended periods of time by increased beta activity in local cortical circuits. We argue however that such an account would make similar predictions to the 'beta-syntax' hypothesis in terms of the beta power modulations that should be observed. Upon encountering a syntactic violation, the demands on working memory should be diminished once the language comprehension system 'realizes' that it is not possible to make the sentence grammatical (the system would essentially give up). This would result in a decrease in beta power (although it may initially increase until the system registers that the sentence cannot be made grammatical). For the OR case however the demands on working memory should increase relative to the SR condition because it is possible to reach a grammatical interpretation of the sentence, and that means more information needs to be maintained online. This in turn should result in an increase in beta power, the opposite of what we observed.

Beta desynchronization in left prefrontal cortex has also been associated with the richness of information encoded in a memory trace, and/or how much information is represented in a memory representation being retrieved (Hanslmayr, Staudigl, & Fellner, 2012). This 'information via desynchronization' hypothesis would suggest that compared to the SR condition, the decreases in beta power we observed for the AVR and OR conditions are related either to the formation of a more information-rich memory representation of the linguistic input up to that point in the sentence, or to the retrieval of more information-rich memory representations after encountering the unexpected target word. The first possibility appears plausible since for both the AVR and OR conditions it is likely that some form of re-analysis is attempted, and so it may be that encoding a richer representation of prior linguistic information after encountering the unexpected target word assists with that re-analysis. The second possibility also appears plausible (and may be in line with cue-based parsing approaches; e.g., Lewis, Vasishth, & Van Dyke, 2006; Van Dyke & Lewis, 2003) because it may be the case that upon encountering the unexpected target word, the language comprehension system retrieves a richer representation of the prior linguistic input, again in order to attempt re-analysis. Both of these possibilities involve a representation of the linguistic context based on the preceding sentence input. In this respect the 'information via desynchronization' hypothesis when applied to sentence comprehension may be closely related to the 'beta-maintenance' hypothesis, with some minor differences in the details. The 'beta-maintenance' hypothesis places more emphasis on network dynamics underlying the representation of sentence-level meaning, while the 'information via desynchronization' hypothesis places more emphasis on the amount of information that can be represented in a network of regions. Since our study was designed to directly compare the 'beta-syntax' hypothesis with the 'beta-maintenance' hypothesis we choose for now to interpret our results as support for the 'beta-maintenance' hypothesis. The extent of overlap between this and the 'information via desynchronization' hypothesis applied to sentence-level language comprehension is well worth investigating in future research.

The timing of our beta effects is in good agreement with previous self-paced reading and eye-movement findings from a study that used Dutch stimuli highly similar to those employed here (Mak et al., 2002). Mak et al. (2002) observed RT differences between SR and OR sentences at the word directly following the relative clause-final auxiliary (target word + 1) in their self-paced reading experiment, and observed differences in first-pass reading times at the auxiliary itself (target word) in their eye-movement experiment. Our time window of interest starts at 540 ms after the onset of the target word, but visual inspection of the TF representations (Figure 5.3A) suggests that the differences in beta power between the AVR and SR and between the OR and SR conditions begins between about 300 and 400 ms after target word onset. These effects extend across the remainder of the time window of interest (until 1200 ms after target word onset), which means they encompass the end of the interval in which the target word is presented, the full interval in which the word directly following the relative clause-final auxiliary is presented (target word + 1), and the beginning of the interval of the word directly following that (target word +2). One might ask why the beta decrease for the AVR condition (a condition that was not present in the study by Mak et al., 2002) also extends across this entire time interval, since after encountering a syntactic violation the sentence can no longer be made grammatical. We suggest it is likely that the language comprehension system attempts some form of re-analysis in the AVR condition as well, but does so unsuccessfully. The beta decrease here would reflect preparation for changes in the configuration of the cortical network responsible for representing the sentence-level meaning, in anticipation of that meaning being updated during re-analysis (even though it is never actually successfully updated).

Finally, both effects appear at left frontal and left temporal sensors (Figure 5.3B), strongly suggesting the involvement of the left hemisphere core language network (left inferior frontal cortex and left temporal cortex; e.g., Hagoort, 2013; Hickok & Poeppel, 2007; Makuuchi & Friederici, 2013). The effect appears to be more widespread (but still restricted to left hemisphere sensors) for the comparison between the SR and OR conditions than for the comparison between the SR and AVR conditions, possibly indicating the recruitment in the OR condition of additional cortical regions that represent new lexical information required for changing the sentence-level meaning representation. These regions would not be recruited to the same extent for the AVR condition because a grammatically acceptable sentence-level meaning representation is never successfully constructed. We acknowledge that this interpretation is speculative, and suggest that a useful next step would be to attempt to localize the sources of these beta effects to obtain a more precise indication of exactly which cortical structures are involved. With improved spatial precision it may be possible to offer a more precise functional interpretation based on prior hemodynamic and other neuroimaging results.

5.4.3 Conclusions

This study directly compares the predictions of the 'beta syntax' and the 'beta maintenance' hypotheses about the role of beta oscillatory activity in sentence-level language comprehension. It shows that MEG beta power over left frontal and left temporal sensors decreases, both when the relative clause-final auxiliary indicates an object-relative reading of the sentence, and when it fails to agree in grammatical person with either the antecedent noun phrase in the matrix clause or the relative clause-internal noun phrase, resulting in an ungrammatical sentence. There is no such beta power decrease when the auxiliary indicates a subject-relative reading of the sentence. This provides compelling support in favour of the 'beta maintenance' hypothesis, and against the 'beta-syntax' hypothesis. In addition to the time-frequency results, we also observed a magnetic P600 effect at the relative clause-final auxiliary when comparing subject-relative clause sentences with both agreement violation sentences and object-relative clause sentences. To the best of our knowledge this is the first report of a P600 effect using MEG for the processing difficulties associated with the comparison between subject- and object-relative clause sentences. Our findings suggest that while both left and right hemisphere cortical sources are additionally recruited to try to resolve the processing difficulties associated with agreement violation sentences, only left hemisphere sources appear to be additionally recruited for resolving the difficulties associated with object-relative clause sentences.

Notes

1. This jittered approach to the presentation of each word was used to minimize the effect of onset- and offset-related evoked activity (with jittered timing such short-lived evoked activity is likely to wash out in the average) on the TF representations of the data.

Chapter 6

Summary and General Discussion

Partially based on:

Lewis, A. G., Wang, L., & Bastiaansen, M. (2015). Fast oscillatory dynamics during language comprehension: Unification versus maintenance and prediction? *Brain and language*, *148*, 51-63.

Lewis, A. G., & Bastiaansen, M. (2015). A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex*, *68*, 155-168.

Lewis, A. G., Schoffelen, J. M., Schriefers, H., & Bastiaansen, M. (2016). A predictive coding perspective on beta oscillations during sentence-level language comprehension. *Frontiers in human neuroscience*, *10*.

Comprehending language requires the intricate interplay between various types of information represented at multiple hierarchical levels (e.g., phonological, semantic, or syntactic information, c.f., Jackendoff, 2007). This in turn requires the intricate interplay between multiple brain regions at different hierarchical levels, forming networks responsible for representing and processing various types of linguistic and non-linguistic information (c.f., Hagoort, 2013, 2014; Hickok & Poeppel, 2007; Makuuchi & Friederici, 2013). In this thesis I investigated neural oscillations as a direct window onto the dynamic coupling and uncoupling of such networks of brain regions supporting language comprehension. Specifically, I explored various aspects of the proposed link between beta oscillations and syntactic processing. Furthermore, I investigated the extent to which oscillatory activity in the beta frequency range (13 to 30 Hz) provides an index of syntactic unification/integration (the 'beta-syntax' hypothesis; Bastiaansen & Hagoort, 2015), or whether it is rather related to the maintenance/change of NeuroCognitive Networks (NCNs) responsible for representing the current sentence-level meaning under construction (the 'beta=maintenance hypothesis'; Lewis & Bastiaansen, 2015; Lewis, Wang, & Bastiaansen, 2015). In this final chapter I will discuss these two proposed roles for beta oscillations during language comprehension in light of the empirical results presented in *Chapters 2* to *5*, and speculate about some potential directions future empirical work in this area might take.

As a brief reminder, both the (*strong* version of the) 'beta-syntax' hypothesis and the 'beta-maintenance' hypothesis propose that more demanding syntactic processing should result in higher beta power. The 'beta-maintenance' hypothesis claims that this is also the case when other types of processing (e.g., semantic processing) become more demanding. Similarly, both the 'beta-syntax' hypothesis and the 'beta-maintenance' hypothesis propose that when syntactic processing is disrupted, and it is clear to the system that the grammaticality of the sentence cannot be recovered, this should result in lower beta power. Again, the 'beta-maintenance' hypothesis claims that this is also the case for disruptions of other types of processing (e.g., semantic processing). Importantly, the 'beta-syntax' hypothesis proposes that when syntactic processing is temporarily disrupted (e.g., when there is ambiguity between alternative syntactic constructions and the linguistic input disambiguates towards the less preferred alternative) but the grammaticality of the sentence can still be recovered, beta power should be ***higher*** due to syntactic processing becoming more demanding after the disruption (syntactic structure building does not halt after disambiguation toward a less preferred syntactic construction at locally ambiguous regions of a sentence; e.g., Frazier & Rayner, 1982). Under the same circumstances, the 'beta-maintenance' hypothesis proposes that this disruption is taken as a cue that the representation of the underlying sentence-level meaning needs to change, and beta power should therefore ***decrease***.

6.1 Beyond single sentences

There is now a sizeable body of literature investigating neural oscillations in relation to the processing of words in isolation (i.e., without a sentence context, e.g., Brennan, Lignos, Embick, & Roberts, 2014; Pulvermüller et al., 1996; van Ackeren, Schneider, Musch, & Rueschemeyer, 2014), and also a moderate (but rapidly expanding) body of literature investigating neural oscillations in relation to sentence-level comprehension (e.g., Bastiaansen, Magyari, & Hagoort, 2010; Kielar, Panamsky, Links, & Meltzer, 2015; Peña & Melloni, 2012). The influence of discourse-level information on oscillatory activity related to sentence comprehension, however, has until now not been addressed. In *Chapter 2* discourse-level semantic coherence between sentences comprising short stories was manipulated in an effort to fill this gap in the literature. Most importantly, this experiment showed that beta oscillatory activity is sensitive to discourse-level semantic coherence, exhibiting higher beta power for coherent short stories than for incoherent ones (Figure 2.2). Although beta has been specifically linked to sentence-level syntactic processing (e.g., Bastiaansen et al., 2010), there are also reports of modulations of beta activity related to semantic (e.g., Wang, Jensen et al., 2012) and other (e.g., Luo, Zhang, Feng, & Zhou, 2010) forms of linguistic processing. Whether the modulation of beta power observed in *Chapter 2* is related to syntactic or semantic processing (or possibly both) cannot be unambiguously established solely on the basis of the study presented there. An important next step will be to use discourse-level manipulations specifically targeted at modulations of local sentence-level syntactic (e.g., Koornneef & Sanders, 2013) or semantic (e.g., Nieuwland & Van Berkum, 2006) processing in order to tease these two possibilities apart.

6.2 Linguistic proficiency matters

If beta power does turn out to be an index of syntactic processing during language comprehension, then questions arise about whether or not similar beta effects are present for non-native speakers of a language, whether such effects are comparable to those found for native speakers, and if not then what factors influence observed differences between native speakers and second language learners? *Chapter 3* addressed some of these issues by investigating the effects of processing grammatical gender agreement violations in Dutch on beta oscillatory power. A comparison was made between Dutch native speakers and German late second language learners (still highly proficient) of Dutch. The first two experiments in *Chapter 3* showed that beta power was lower upon encountering a grammatical violation within a sentence for native speakers but not for late second language learners. This suggests that one's level of proficiency matters for whether or not syntax-related beta oscillatory activity is modulated by disruptions of syntactic processing. The third and fourth experiments of that chapter established that late second language learners do show syntax-related beta modulations (again, lower beta power for grammatical violations) when the task demands that they explicitly pay attention to grammatical information. Perhaps more importantly, those experiments showed that this is only the case when trials are re-sorted according to the second language learners´ (often incorrect) subjective lexical representations of grammatical gender information associated with any particular noun. It thus appears that late second language learners do not make use of grammatical gender information in the course of processing determiner-noun pairs in Dutch (or rather, disrupted processing of this information does not modulate beta oscillatory activity), but when explicitly forced to focus on such information they rely on their subjective long-term memory representations of that information.

While these studies all address beta oscillations in relation to syntactic processing, the findings can be explained equally well as related to maintenance/change of the NCN responsible for the representation of the current sentence-level meaning under construction. In all cases the violation acts as a cue to the system that the current sentence-level meaning is incorrect and may need to change (although for grammatical gender violations this sentence-level meaning likely does not change, the system still temporarily anticipates that this will be necessary before registering the sentence as ungrammatical instead). For late second language learners these cues are only used by the system in case grammatical information is highlighted as important for the task, and that's why beta effects are only observed under a limited set of circumstances for this group of participants. Furthermore, the system is only able to use violations as cues indicating (temporarily) that a change in the current sentence-level meaning is necessary, in case those violations are perceived as such. This means that the system will be sensitive to subjectively perceived violations of grammatical gender agreement rather than objectively ungrammatical agreement relations (which may go unnoticed by the language comprehension systems of late second language learners). After all, the only representations that can be relied upon are the ones that are there in the first place, whether they are objectively correct or not.

6.3 Beta relates to maintenance/change

In *Chapter 5* the two hypothesized roles for beta oscillations (syntactic integration/unification versus maintenance/change of a NCN) during sentence-level processing were directly compared. The two hypotheses make opposing predictions for how beta activity should be modulated when processing the disambiguating element in locally ambiguous object-relative clause sentences in Dutch: 1) the (*strong* version of the) 'beta-syntax' hypothesis predicts that beta power should increase because the local ambiguity requires additional processing (syntactic structure building becomes more demanding) in order to recover the correct meaning of the sentence; 2) the 'beta-maintenance' hypothesis predicts that beta power should decrease because the disambiguating element acts as a cue to the comprehension system, indicating that the input does not match the current sentence-level meaning representation up to that point, and that this representation needs to be changed. Crucial to this distinction is that syntactic processing is not disrupted at the point of disambiguation within these locally ambiguous object-relative clause sentences, and the grammaticality of the sentences can be recovered (while for the grammatical violation sentences from *Chapter 5* it is clear to the system at the point of disambiguation that the grammaticality of the sentence cannot be recovered). Furthermore, local ambiguity is the key to the prediction of the 'beta-maintenance' hypothesis, where the comprehension system is assumed to have already (at least partially) committed to the preferred sentence structure (subject-relative clause), and upon encountering the disambiguating element in the sentence indicting that an alternative, less preferred sentence structure is present (object-relative clause), needs to change the current sentence-level meaning accordingly. This is different from cases without local ambiguity (i.e., where it is clear early in the sentence that the sentence is unambiguously an object-relative clause construction), where both hypotheses predict higher beta power compared to subject-relative clause sentences due to higher processing demands. Clear evidence was found in favour of the 'beta-maintenance' hypothesis (beta power decreased for object-relative clause sentences), and against the (*strong* version of the) 'beta-syntax' hypothesis.

It is important to realize that this does not completely exclude a link between beta activity and syntactic processing during sentence comprehension, but rather the beta-syntax link can be subsumed under the more general overarching beta-maintenance hypothesis. Many results have indicated a link between beta activity and syntactic processing, and it seems clear that many syntactic manipulations will result in modulations of beta oscillations. At the same time, there are other (non-syntactic) linguistic manipulations (reviewed in *Chapter 1*) that result in a modulation of beta power. Moreover, *Chapter 5* shows that certain kinds of syntactic manipulations do not necessarily result in a modulation of beta power in the direction predicted by a strict beta-syntax mapping. The link between beta activity and syntactic processing will clearly still prove useful in future investigations. What I have shown in this thesis however, is that one should always keep in mind that what these beta modulations are actually capturing is the active maintenance or change of the current NCN responsible for representing the sentence-level meaning under construction. This may often correspond directly to either ongoing (increased beta activity) or disrupted (decreased beta activity) syntactic integration/unification (i.e., one aspect of the construction of a sentence-level meaning representation), but that is not always going to be the case.

A major lesson to be learned is that when manipulations of syntactic processing during sentence comprehension result in modulations of beta activity, one can be relatively certain of a link between the two. The reverse is however not necessarily true, as modulations of beta activity during sentence comprehension have also been linked to other types of processing (e.g., semantic processing; Wang, Jensen et al., 2012). The relationship is not veridical in this sense, but that does not rule out beta as a useful tool for investigating ongoing syntactic processing under the right set of circumstances.

A number of important questions remain unanswered, and should be addressed by future research. For instance, it would be useful to have a direct measure of the exact brain regions recruited by the language comprehension system to form part of the NCNs that are being actively maintained/changed by beta oscillatory activity. I have made some inferences in this regard in *Chapter 1,* based on neuroimaging methodologies with better spatial resolution than EEG or MEG. However, the use of source reconstruction techniques with electrophysiological data is important in future language comprehension studies in order to gain more fine-grained insights into the spatial distribution of the cortical networks whose temporal dynamics are being investigated.

The reader may have noticed that despite claims to the effect that superior temporal information can be obtained using EEG and/or MEG, relatively little discussion has been devoted to the precise timing of the various beta oscillatory effects reported in this thesis, and indeed the same is true of the wider literature on oscillatory phenomena related to sentence comprehension. It will be important for future research to pay closer attention to this aspect of oscillatory data in relation to sentence comprehension. The benefit will be for instance to disentangle oscillatory phenomena related to early (perhaps more automatic) from those related to late (perhaps more controlled) stages of linguistic processing (see e.g., Brouwer, Fitz, & Hoeks, 2012 for a similar discussion in the ERP literature).

It may turn out that the oscillatory phenomena we are investigating are initially modulated in one direction (e.g., early decrease in power relative to the event of interest), and at later time points are modulated in the opposite direction (perhaps showing rebound effects with late power increases) for oscillatory activity in the same frequency range. Such opposing effects may reflect different aspects of cognitive processing, and this may be especially likely for more rapid oscillatory phenomena in the beta and gamma frequency ranges. These possibilities remain speculative for now, but they highlight some potentially interesting directions for future work if we start paying closer attention to the precise timing of oscillatory phenomena measured during sentence comprehension.

6.4 Predicting the future

An interesting question that has come to the fore of late is the extent to which predictive processing plays a crucial role in language comprehension (e.g., Dell & Chang, 2014; Huettig, Mani, & Huettig, 2015; Huettig, 2015; Kuperberg & Jaeger, 2015; Pickering & Garrod, 2014). In *Chapter 4* an attempt was made to develop a measurement tool using oscillatory EEG entrainment and memory reactivation effects to track exactly when during sentence comprehension lexical information becomes activated. The idea was to attach so-called 'frequency tags' (Wimber, Maaß, Staudigl, Richardson-Klavehn, & Hanslmayr, 2012) to lexical items learned in a study phase, and to track any frequency-specific reactivation of these 'frequency tags' during reading. This would facilitate the investigation of whether or not there was any information in the EEG signal about the activation of a particular lexical item (or of particular lexical information) with which the 'frequency tags' were associated, at various positions within a sentence. The hope was that this would provide a useful tool for investigating predictive processing in the form of lexical pre-activation. Unfortunately, the replication experiment reported in *Chapter 4* produced only mixed results, and there we argued that without more robust memory reactivation findings this methodology could not be used in the way originally envisaged. The approach itself has a range of potential uses (even beyond the investigation of predictive processing), and future work should try to further refine the procedure to try to make it more robust, so that it might be confidently implemented to track the (pre-) activation of lexical information.

A currently popular group of explanations for how predictive processing in the form of Bayesian hierarchical inference might be implemented in the brain are predictive coding theories (e.g., Clark, 2013; Friston, 2005). In the remainder of this discussion I will outline how the literature linking beta oscillations to sentence comprehension might be captured by one such predictive coding framework (Friston, 2005). There are still many details to be worked out, but the basic outline may provide an excellent stepping-stone to future investigations into the link between beta oscillatory activity and sentence comprehension.

6.4.1 Language comprehension, neural oscillations, and predictive coding

It has been shown experimentally in monkey visual cortex and in rat somatosensory cortex that gamma oscillations are most prominently expressed in supragranular cortical layers (L2/3), while beta oscillations are more prominent in infragranular (and granular) layers (L4/5; Maier, Adams, Aura, & Leopold, 2010; Roopun et al., 2006, 2008). At the same time, feedforward connections predominantly originate in superficial layers (L2/3) and terminate in L4, while feedback connections originate from deeper layers (L4/5) and terminate outside of L4 (Bastos et al., 2012). This has led to the proposal (Wang, 2010) that within cortical hierarchies, feedforward signaling may be mediated by high frequency oscillations (in the gamma range for instance) compared to feedback signaling, which may be mediated by oscillations at lower frequencies (in the beta or alpha range). Bastos et al. (2012) have suggested that this principle might constitute a canonical form of hierarchical functional organization in the brain. The proposal is that within a cortical processing hierarchy gamma oscillations might predominate for bottom-up interactions, while beta oscillations might predominate in the top-down direction. The levels of such processing hierarchies can be restricted to local cortical regions (e.g., occipital cortex for most of the visual system), but can also span non-local cortical regions (e.g., left inferior frontal cortex and left middle temporal cortex for two important parts of the core language processing hierarchy).

Bastos et al. (2012) have also proposed that this canonical hierarchical organizing principle might provide physiological correlates of the implementation of predictive coding within cortical hierarchies. From a predictive coding perspective, top-down information (conveyed by feedback connections) provides context for lower-level processing over slower time scales (beta oscillations), while bottom-up information (conveyed by feedforward connections) works on faster time scales (gamma oscillations) propagating prediction errors up the hierarchy in order to rapidly adapt predictions at these higher levels. This implies that oscillatory activity in the beta frequency range might be a proxy for top-down predictions about activity at lower hierarchical levels within a cortical hierarchy, while gamma oscillations might be an indication of the forward propagation of prediction errors to higher cortical levels in order to update predictions (Bastos et al., 2012; Friston, Bastos, Pinotsis, & Litvak, 2015).

I have argued that the construction of a sentence-level meaning representation during unification entails the formation of a NeuroCognitive Network (NCN), encompassing areas in left inferior frontal cortex, left temporal cortex, and left inferior parietal cortex (Hagoort, 2013, 2014), along with other relevant areas outside the core language network, depending on the particular context in which sentence-level meaning construction is taking place (e.g., recruitment of the theory of mind network for taking another person's perspective). By 'sentence-level meaning' I am referring not just to the semantics associated with the individual words comprising a sentence, but also to the semantics derived from the syntactic structure governing the hierarchical relations between those words. As pointed out by Bressler and Richter (2015), beta oscillations could be simultaneously involved in both the maintenance of the current NCN, as well as the propagation of top-down predictions (perhaps based on the information represented or processed in the NCN) to lower levels in the cortical processing hierarchy (Figure 6.1). I would like to emphasize that these two related roles for beta oscillations are entirely compatible with one another. The representation of information in a distributed NCN and the use of that information to make predictions in a top-down fashion constitute closely related and heavily interdependent forms of neural processing. In the case of sentence-level language comprehension, beta synchrony may therefore be responsible for the active maintenance of the current NCN supporting sentence-level meaning construction, as well as the top-down transfer of predictions that the sentence-level meaning might convey to lower levels (e.g., the memory component responsible for lexical retrieval) of the cortical processing hierarchy. Such predictions can be about individual words, but also about other units of linguistic information (e.g., particular syntactic constructions; cf. Levy, 2011).

Bastos et al. (2012) have also discussed the role of gamma oscillations in the computation of prediction errors and the passing forward of those prediction errors to hierarchically higher cortical levels in order to update the generative models (and hence predictions) at those levels. There are arguments linking gamma to prediction errors in sentence comprehension (c.f., Lewis & Bastiaansen, 2015), but since this thesis is about beta oscillations those ideas will not be discussed any further here.

Simplified illustration of hierarchical information flow during language comprehension

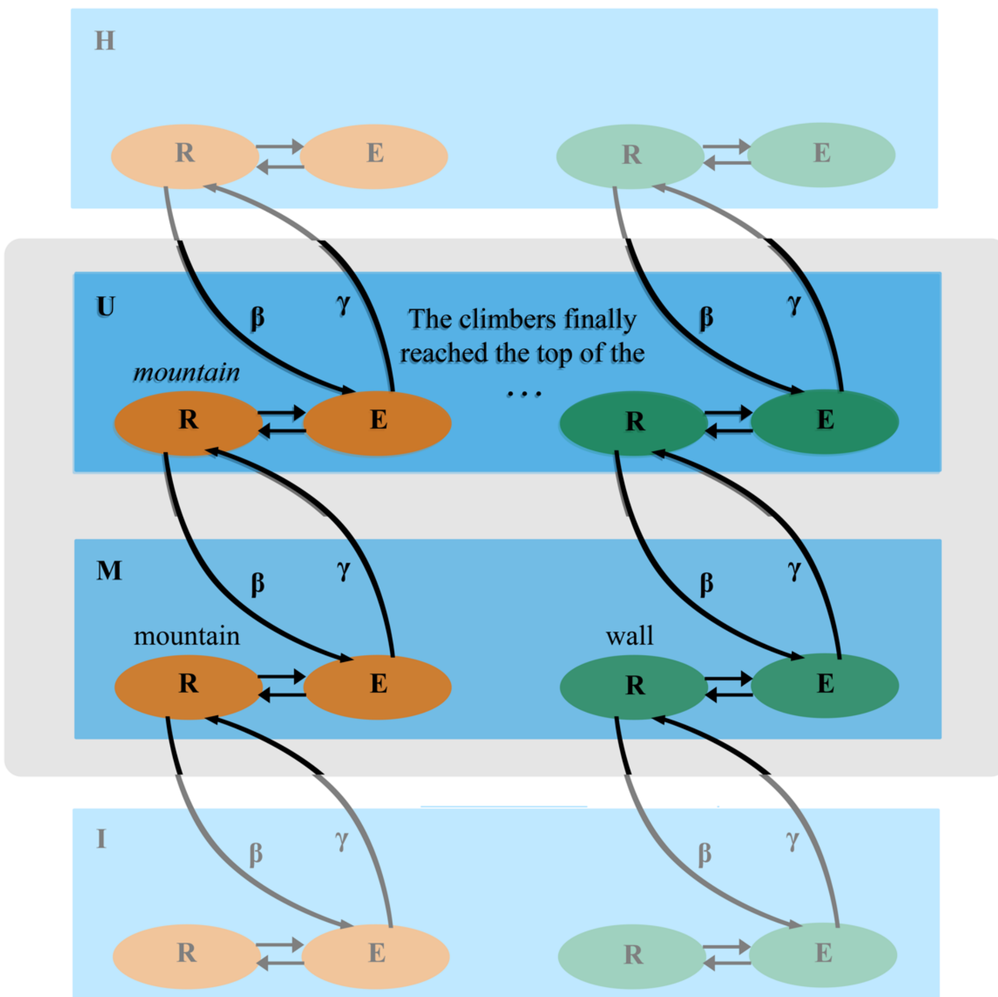

Figure 6.1 Simplified illustration of the proposed hierarchical flow of information during language comprehension. Blue boxes refer to different levels of the processing hierarchy. We focus (highlighted portion of figure) on levels corresponding to the memory component (M) and the unification component (U) from the MUC framework (Hagoort, 2013). Also pictured (but not highlighted) are an input level (I) for auditory, visual or other types of input to the language comprehension system, and a higher level (H)

for things like cognitive control or other forms of higher-level processing. Within each hierarchical layer there are multiple representation (R) and error (E) units. Those in orange are relevant for our example, while those in green indicate (more than one pair of) potential competing representation and error units. Not pictured in the figure (to avoid the figure becoming too cluttered) are inhibitory lateral connections between these (orange and green) representation and error units at the same hierarchical level, responsible for lateral inhibition of competing representations. In the example from the figure, a reader has read the input 'The climbers finally reached the top of the' and makes a strong prediction that the next word will be '*mountain*'. We have suggested that nodes in the NeuroCognitive network responsible for constructing and maintaining a representation of the sentence-level meaning are linked via oscillatory activity in the beta frequency range. This is not depicted in the figure but the network comprises (amongst others) nodes from left inferior frontal gyrus, left inferior parietal and left temporal cortex. Based on this sentence-level meaning, the system generates a prediction at U. This prediction is sent down the processing hierarchy (via oscillatory activity in the beta band) from an R unit at U, to an E unit at M, where the prediction is matched with incoming linguistic information (sent from the R unit at M to the E unit at that same level) to compute a prediction error. That prediction error is sent back up the processing hierarchy (via oscillatory activity in the high gamma range) from an E unit at M to an R unit at U (but also to an R unit at M to update representation units at the same level) so that the generative model (and hence the prediction) at this higher level can be updated if necessary. When the input matches a strong prediction (e.g., in our example the input is 'mountain') this results in an increase in low-mid gamma power reflecting the match, as well as strong lateral inhibition (not depicted in the figure) of competing representation and error units at the same hierarchical level.

6.4.2 Beta oscillations and top-down predictions

The studies discussed in this section have already been introduced (initially in *Chapter 1*) in the context of the proposed beta-syntax link and/or the proposed beta-maintenance link. Next I discuss these studies in relation to the predictive coding framework just described.

A number of studies report that beta power is sensitive to both syntactic violations (Bastiaansen et al., 2010; Davidson & Indefrey, 2007; Kielar, Meltzer, Moreno, Alain, & Bialystok, 2014; Kielar et al., 2015; Pérez, Molinaro, Mancini, Barraza, & Carreiras, 2012) and semantic incongruities (Kielar et al., 2014, 2015; Luo et al., 2010; Wang, Jensen et al., 2012). In all of these studies, beta power was higher following some target word for syntactically and semantically acceptable sentences compared to target words that resulted in a syntactic violation or a semantic incongruity. Similarly, Luo et al. (2010) showed that beta power was higher for rhythmically normal compared to abnormal target nouns in Chinese verb-noun pairs. In addition to grammatical violations, Pérez et al. (2012) showed that beta power following a target word was lower for the case of Spanish 'Unagreement' (where the sentence remains grammatical despite a mismatch between the grammatical person feature marking on the subject and that on the verb of a sentence) compared to grammatically acceptable target words. These studies all have in common that there is some cue in the linguistic input (e.g. syntactic violation, semantic incongruity, etc.)

that (temporarily) indicates to the language comprehension system that the current representation of the sentence-level meaning needs to change (even if it does not actually change in the case of grammatical violations for instance). This may result in the system assigning less value to top-down predictions (based on the current sentence-level meaning) in these cases, as that information has proven unreliable, and under the predictive coding framework outlined above this would result in the observed decrease in beta activity. The findings from *Chapters 3* and *5* can be given a similar explanation. In both those chapters either syntactic violations, or in the case of the object-relative clause sentences in *Chapter 5* the disambiguating auxiliary indicating that the less expected syntactic structure was to be preferred, would constitute cues to the language comprehension system that the current sentence-level meaning needs to change. In all these cases the expected beta power decrease was observed, potentially signaling less reliance on top-down predictions during comprehension.

Another group of studies has shown that beta activity is higher when (fully grammatical) sentences are more syntactically demanding, without requiring a change of the sentence-level meaning representation (Bastiaansen & Hagoort, 2006; Meyer, Obleser, & Friederici, 2013; Weiss et al., 2005). Bastiaansen and Hagoort (2006) reported that beta power was higher for syntactically more demanding center-embedded compared to right-branching relative clauses. Meyer et al. (2013) showed that beta power was higher for long- compared to short-distance subject-verb agreement dependencies at the point in the sentences where the dependency could be resolved. Weiss et al. (2005) found higher beta coherence between frontal and posterior electrode sites for syntactically more complex object-relative compared to subject-relative clauses[1]. In all of the above cases the beta increase may indicate a greater reliance on top-down predictions based on the current sentence-level meaning (i.e., the increased activity may be related to greater weighting of the top-down signal based on the current generative model), in order to actively try to integrate the new linguistic input into the current sentence-level meaning representation.

It is important to notice that these studies differ from the study in *Chapter 5* comparing subject- and object-relative clause sentences, in that for the studies reported above (unlike in *Chapter 5*) there is no clear cue to the language comprehension system indicating that the current sentence-level meaning needs to change. Instead, when it encounters the point in the sentence where syntactic unification load increases, the system may take this as a cue that the current sentence-level meaning needs to be actively maintained.

Bastiaansen et al. (2010) showed that beta power increased linearly over the course of syntactically legal sentences, but returned to baseline levels at the point of a syntactic violation within sentences. They also showed that lists of the same words contained in the sentences in random order (no syntactic structure) did not exhibit any increase in beta power over the course of presentation of the lists (see also Bastiaansen & Hagoort, 2015). The beta decrease at the point of a syntactic violation can be explained as before, where the violation acts as a cue to the language comprehension system indicating that the current sentence-level meaning should not be relied upon as heavily as before, resulting in a lower weighting of top-down predictions and a concomitant decrease in beta activity. For random word lists no sentence-level meaning is constructed, and hence no beta power increase over the course of presentation of the word list, indexing reliance on top-down predictions based on that sentence-level meaning, should be expected.

At this point I would like to reiterate that a role for beta activity in the transfer of top-down predictions within cortical hierarchies is entirely compatible with a role for beta activity in the maintenance/change of NCNs representing the current sentence-level meaning under construction (c.f., Bressler & Richter, 2014). When the language comprehension system actively maintains the current sentence-level meaning (and underlying NCN) top-down information will be relied upon more heavily, both of which should lead to increased beta activity. Similarly, when the system encounters a cue indicating that the current sentence-level meaning needs to change (along with the underlying NCN) top-down information will be considered less reliable by the system, again both leading to decreased beta activity. These two complementary roles for beta oscillatory activity during sentence comprehension offer exciting new possibilities for investigating neural dynamics during language comprehension.

6.5 Concluding remarks

To conclude, this thesis shows that discourse-level information, as well as linguistic proficiency with a particular language, are two important factors that affect beta band oscillatory neural activity during comprehension. More importantly, it shows that the strict link between beta band oscillatory activity and syntactic integration/unification proposed by the 'frequency-based segregation of syntactic and semantic unification hypothesis' does not always hold. Instead, beta activity during language comprehension appears to be related to the maintenance/change of the NCN responsible for representing the current sentence-level meaning under construction.

There are numerous unanswered questions, and I have briefly outlined a few exciting new directions that might be explored in future research linking beta oscillatory activity to sentence-level language comprehension, particularly with regard to prediction and predictive coding. The conclusion of this thesis therefore provides an excellent starting point from which new investigations can be launched into the relationship between beta oscillatory activity and sentence comprehension. These investigations can now take as their starting point the fresh perspective that beta is related to the maintenance/change of a NCN representing the sentence-level meaning under construction, and speculatively also to top-down predictions based on that sentence-level meaning. This line of research has the potential to uncover the secrets of how different brain regions interact rapidly and dynamically in order to allow different types of linguistic and non-linguistic information to do the same, such that readers/listeners succeed in correctly interpreting the message their interlocutor intends to convey. The future of this research therefore appears bright, or at the very least, a whole lot clearer.

Notes

1. Importantly, the constructions used by Weiss et al. (2005) are not locally ambiguous (in contrast to the subject vs object relative clauses of chapter 5), and so it is clear from the beginning of the relative clause whether it will be a subject or an object-relative construction. This means that the main difference between these conditions is in the syntactic complexity across the relative clause. There is no point within the relative clause at which an a priori preferred structure (e.g. subject-relative) is replaced by an alternative, less frequent structure (e.g., object-relative).

Appendices

Appendix A: Chapter 2 Stimuli

Coherent stories:

1 Charles verliet zijn vaderland Senegal om in Europa te werken.
Met een levensgevaarlijk klein bootje werd hij naar Tenerife gesmokkeld.
Hij moest daar hard werken voor een klein beetje geld.
Zijn familie had het geld dat hij stuurde hard nodig.

2 Een groep hangjongeren maakte zaterdag enorme rotzooi in het parkje.
Twee van hen trapten die avond een bushokje aan diggelen.
Toevallig kwam een agent de hoek om die hen arresteerde.
De kwajongens kregen driehonderd uur taakstraf bij de plantsoenendienst opgelegd.

3 De familie Witteveen verbleef in hun zomerhuis in het woud.
Een avond hadden ze een taart achtergelaten in de keuken.
Die nacht sloop een beer binnen die de taart opvrat.
Pas de volgende ochtend ontdekte de geschrokken vader de inbraak.

4 Die gladde autoverkoper probeerde Lieneke eens een auto te verkopen.
De wagen zag erg goed uit maar was te duur.
Twee weken later bleek haar buurman hem te hebben gekocht.
Maar na een jaar moest hij al naar de sloop.

5 De reggaeartiest was geboren in een gewelddadige stad op Jamaica.
Hij had veel agressie meegemaakt en hij blowde uitzonderlijk veel.
Dit inspireerde hem bij het schrijven van zijn hoopvolle teksten.
Zo werd hij op Jamaica populair en brak wereldwijd door.

6 Net op tijd stopte Jaaps auto voor het rode licht.
In het Peugeootje naast hem zag hij een leuke meid.
Hij reed weg bij groen en zij ging er achteraan.
Thuis herkende hij haar ineens, ze was zijn nieuwe buurvrouw.

7 De sterke schildknaap Karel diende ridder Archibald met onuitputtelijke kracht.
Tijdens de veldslag verpletterde Karel vele vijanden met zijn schild.
Daarom was ridder Archibald zeer tevreden over de dappere jongeman.
Met instemming van de koning werd hij tot ridder geslagen.

8 Bij de Rietveldse Weide was een schaap te water geraakt.
Het arme beest kon niet meer uit de sloot komen.
's Avonds wandelde de oude kromme boer toevallig langs de weide.
Hij alarmeerde twee dorpsbewoners en kon zo zijn dier redden.

9 Matthias was net uitwonend en vierde dat in een café.
Midden in de nacht fietste hij naar zijn nieuwe huis.
Hij reed door rood en z'n lichten waren allebei kapot.
De politieagent bij het stoplicht zag het door de vingers.

10 Piet en Jeremy gebruikten veel drugs tijdens hun studie rechten.
Op vakantie in Zuid-Amerika probeerden ze cocaïne mee te smokkelen.
Bij de Nederlandse douane werden ze door een drugshond gesnapt.
Door hun strafblad nam geen enkel advocatenkantoor hen later aan.

11 Deze bekende Franse schilder trok honderdvijftig jaar geleden door Frankrijk.
Zijn schilderstijl werd nog niet goed begrepen in zijn tijd.
Hij kon nauwelijks rondkomen en bedelde zijn bestaan bij elkaar.
Nu betalen rijkelui vele miljoenen euro's voor zijn kleinste werkjes.

12 Mathilde zat in de bioscoop een romantische film te kijken.
De film ging over een onmogelijke verhouding tussen twee geliefden.
De twee waren lid van verschillende rivaliserende en gewelddadige jeugdbendes.
Toen de jongen werd doodgeschoten hield Mathilde het niet droog.

13 Al de hele zomer droomde Klaas van schaatsen op natuurijs.
In januari vroor het voor het eerst sinds jaren hard.
Als eerste schaatste hij op het ijs maar zakte erdoor.
Verkleumd kwam hij terug maar was toch voldaan en blij.

14 's Avonds stapte een agressieve jongen op de bus richting centrum.
Hij bedreigde de chauffeur en rende weg met zijn geldkistje.
De chauffeur ging achter hem aan met een dikke honkbalknuppel.
Hij sloeg hem op zijn hoofd en belde de politie.
15 Irene en Marloes waren sinds de basisschool altijd hartsvriendinnen geweest.
Irene kreeg op haar dertiende een vier jaar oudere vriend.
Marloes werd stikjaloers en bazuinde overal rond dat hij vreemdging.
Maar hij was echter inderdaad vreemdgegaan en de vriendschap herstelde.
16 Odysseus verliet zijn jonge vrouw om te vechten tegen Troje.
Zijn beeldschone Penelope bleef trouw wachten ondanks de vele huwelijksaanzoeken.
Na twintig lange jaren rondzwerven kwam Odysseus terug op Ithaka.
Maar Penelope herkende hem niet meer bij hun eerste ontmoeting.
17 Marie-Louise was verlaten door haar partner en stortte helemaal in.
Ze zocht hulp en praatte over haar eenzame droeve leven.
De therapeut was zelf ook eenzaam en kreeg bijzondere gevoelens.
Marie-Louise fleurde weer helemaal op terwijl de therapeut depressief achterbleef.
18 In het jaar 1953 vond in Zeeland de watersnoodramp plaats.
Wereldwijd werd via de media over de rampzalige gevolgen bericht.
Uit alle windstreken werden heel veel hulpgoederen naar Nederland gestuurd.
Zoveel zelfs dat een gedeelte moest worden teruggestuurd of vernietigd.
19 Een Eskimo trok 's winters door het hoge noorden van Groenland.
Met een hondenslee was hij op zoek gegaan naar voedsel.
's Nachts begonnen zijn honden te blaffen voor een ijsbeer.
Hij doodde het dier en keerde terug naar zijn familie.
20 Familie Beerenbrouck kreeg zeer belangrijke gasten te eten op zaterdagavond.
De gastvrouw schrok toen haar gasten graag rosÃ© wilden drinken.
Zenuwachtig mengde ze rode en witte wijn in een karaf.
Tot haar vreugde had het mengsel een bijzonder lekkere afdronk.
21 Johnny werkte in de jaren zestig bij de DAF autofabriek.
Hij droomde van een eigen auto maar had onvoldoende geld.
Daarom nam hij stiekem iedere dag een los onderdeel mee.
Zo verzamelde hij in vijf jaar een autootje bij elkaar.
22 Billy the Kid was een jeugdcrimineel uit het Wilde Westen.
Al op zijn twaalfde begon hij met diefstal en moord.
Hij ontsnapte een paar keer uit gevangenissen en werd achtervolgd.
De sheriff doodde hem toen hij nietsvermoedend een saloon binnenstapte.
23 Khaldun was een belangrijk Arabische geleerde uit de veertiende eeuw.
Hij was historicus en econoom en grondlegger van de sociologie.
Wetenschap werd in West-Europa destijds geblokkeerd door de katholieke kerk.
Maar in de Arabische wereld kon hij vrijuit wetenschap bedrijven.
24 Hans Brinker was het jongetje dat een belangrijke dijk redde.
's Avonds liep hij eens in z'n eentje langs de dijk.
Halverwege zag hij opeens een klein gaatje waaruit water stroomde.
Hij stak zijn vinger erin en voorkwam zo een dijkbreuk.
25 De bankdirecteur wilde een gouden ketting voor zijn vrouw kopen.
Hij ging naar een hele dure juwelier in de stad.
Zijn vrouw werd vierentwintig jaar en hij wilde haar verrassen.
Helaas moest ze bekennen dat ze de ketting al had.
26 Walter had weinig tijd om zijn onderzoek af te ronden.
Hij kon die zomer nauwelijks proefpersonen vinden voor zijn MRI-experiment.
Maar in zijn kostbare MRI data vond hij goede resultaten.
Zo kon hij toch publiceren en beginnen aan nieuw onderzoek.
27 Het gezin ElQasr ging op vakantie naar familie in Marokko.
Ze pasten met het hele gezin net in de auto.
De wagen was tot op het dak volgeladen met bagage.
Maar na twintig uur rijden ging de auto helaas stuk.

28 Falco was tot levenslang veroordeeld voor een aantal gruwelijke moorden.
Op een avond probeerde hij via aaneengeknoopte lakens te ontsnappen.
Hij klauterde uit het raam maar hoorde een scheurend geluid.
Hij viel schreeuwend naar beneden en was op slag dood.
29 De muren van de studeerkamer waren bedekt met dikke boeken.
Op de werktafel lagen ook nog torenhoge stapels papieren uitgespreid.
De professor liep bedachtzaam naar een van de vele boekenkasten.
Zonder aarzelen vonden de handen van de geleerde een boek.
30 Elke morgen kwam de jonge melkboer vrolijk door de straat.
Alle mensen waar hij melk bracht waren blij met hem.
Ze nodigden hem vaak uit voor een lekker bakje koffie.
Hij was 's avonds vaak alleen en genoot van hun gezelschap.
31 Door de kredietcrisis ging de bank waar Jean werkte failliet.
Drie weken lang was hij van slag door het faillissement.
Maar hij knapte op en besloot een boerenbedrijf te beginnen.
Na een moeilijke start maakte het bedrijf al snel winst.
32 De elf spelers liepen vol zelfvertrouwen het uitverkochte stadion in.
Toen de beroemde spits het veld betrad juichte het publiek.
Door de luidsprekers schalden de klanken van het nationale volkslied.
Iedereen was er van overtuigd dat het elftal zou winnen.
33 Nicolaas was bakker in het kleine zaakje op het plein.
Meestal moest hij 's nachts werken en dan sliep hij overdag.
Daardoor had hij weinig tijd voor z'n vrouw en kinderen.
Maar hij bracht altijd vers brood mee voor het ontbijt.
34 Deze populaire Russische journalist schreef kritische artikelen over de regering.
Hij kritiseerde de oorlog in TsjetsjeniÃ« en was tegen Poetin.
Ook nam hij geregeld deel aan demonstraties voor meer persvrijheid.
Op een ochtend werd hij tijdens zijn dagelijkse wandeling doodgeschoten.
35 Het laatste proefwerk Engels van de brugklas was moeilijk geweest.
De scholieren waren dus allen zeer benieuwd naar de cijfers.
De cijfers van het tentamen waren gelukkig beter dan verwacht.
De leraar Engels zag dit als bevestiging van zijn onderwijsmethode.
36 De nieuwe film kreeg goede recensies in de nationale pers.
De film trok dus heel veel publiek naar de bioscoop.
In de grote bioscopen waren alle voorstellingen bijna helemaal uitverkocht.
Ook in de Nijmeegse bioscopen moest men kaartjes vooraf reserveren.
37 Román was een zigeuner en woonde vijf kilometer van Boedapest.
Hij woonde in een armoedige woonwagen en was heel muzikaal.
Daardoor werd hij een veel gevraagd violist en verdiende goed.
Maar ondanks het geld bleef hij in zijn woonwagen wonen.
38 Thierry was een heel serieuze wielrenner die meedeed aan de Tour.
Hij trainde uitzonderlijk goed en moest niets van doping hebben.
Sommige van zijn collega's waren daar echter wat makkelijker in.
Hij bleef lange tijd fietsen maar won nooit de Tour.
39 Het gerenommeerde architectenbureau mocht eindelijk het nieuwe stadhuis gaan ontwerpen.
Het zou aan de rand van de grote markt komen.
Maar rond de grote markt stonden uitsluitend hele oude gebouwen.
Het stadhuis moest dus zorgvuldig aan de stijl worden aangepast.
40 Het restaurant aan het kerkplein stond bekend om zijn vissoep.
De vissoep werd dagelijks vers bereid door de chefkok zelf.
Maar vandaag was de chefkok door ziekte helaas niet beschikbaar.
De eigenaar overwoog het restaurant daarom helemaal niet te openen.
41 Pietje en Klaasje waren stiekem met vuur aan het spelen.
Ze hadden in de schuur de barbecue met petrolie aangestoken.
De vlammen werden gauw hoog en waren niet te doven.
De brandweer moest toen uitrukken om het vuur te blussen.

42 Klein Duimpje werd door zijn stiefmoeder naar het bos gestuurd.
Om de weg terug te vinden strooide hij kleine steentjes.
Maar zijn stiefmoeder zorgde dat hij geen steentjes meer had.
Tevergeefs probeerde hij later een spoor van kruimeltjes te maken.

43 De ambtenaar staarde uit zijn raam op het saaie gemeentehuis.
Bijna was het vijf uur en mocht hij naar huis.
Tien minuten later liep hij met zijn koffertje naar huis.
Om zes uur at hij thuis zuchtend aardappelen met groente.

44 Afgelopen vrijdag ging dikke tante Joke zoals altijd boodschappen doen.
Onderweg bleef ze voor de etalage van de banketbakker staan.
Ze twijfelde maar kon de moorkoppen helaas toch niet weerstaan.
Zaterdag vertelde ze met veel drama dat ze ging lijnen.

45 In de uitgestrekte bossen in Amerika woonde een sterke houthakker.
Met een zware bijl hakte hij dagelijks hele bomen om.
Met een bevriende houthakker zaagde hij planken van het hout.
Daarvan had hij een blokhut gebouwd waarin hij nu woonde.

46 Gino was een ouderwetse ijscoman met een klein grappig karretje.
De wagen was net koel genoeg voor de Nederlandse zomers.
Maar op een dag in juli was het uitzonderlijk heet.
Toen besloot hij het smeltende ijs als milkshakes te verkopen.

47 Kassameisjes Mien en Pien zaten te tantebetten in de trein.
Een keurige heer van middelbare leeftijd ergerde zich aan hen.
De heer kuchte steeds duidelijker maar kreeg geen telkens respons.
Plotseling viel hij gigantisch uit waarop de meisjes geschrokken wegvluchtten.

48 Een verstrooid omaatje liep met haar winkelwagentje door de supermarkt.
Ze zocht blikjes tomatensoep die ze maar niet kon vinden.
Toen stootte ze per ongeluk een grote stapel blikjes om.
Tot haar verassing bleken het de blikjes tomatensoep te zijn.

49 Chagrijnige Harry ging met zijn vrouw op vakantie per auto.
Hij zat achter het stuur en zijn vrouw las kaart.
In Frankrijk reden ze verkeerd en hij begon te schelden.
Toen zag ze dat zijn kaart vijfentwintig jaar oud was.

50 Een klassiek orkest gaf altijd spetterende optredens in de schouwburg.
De dirigent van dat orkest was heel bekend in kunstkringen.
Hij was een erg ijdele maar ook zeer gepassioneerde man.
Veel vrouwen uit het orkest hadden een zwak voor hem.

51 In de zomer ging boer Piet zijn gras weer maaien.
Op droge dagen hooide hij het gras met z'n trekker.
Toen het gras gedroogd was draaide hij er hooibalen van.
Zo had hij wat extra voedsel voor z'n vee 's winters.

52 Maandagochtend had Michel grote haast omdat hij een tentamen had.
Hij had zich verslapen en fietste razendsnel door de stad.
Maar bij de tentamenzaal was bij aankomst nog helemaal niemand.
Hij was vergeten dat de wintertijd was ingegaan die zondag.

53 Vroeger reisde men per trekschuit tussen de steden in Holland.
Dit was een soort boot die getrokken werd door paarden.
Je voer over het rustige water van sloten en kanalen.
Maar het duurde uren voordat je je bestemming bereikt had.

54 De jonge prinses liep vol verlangen rond de grote vijver.
Ze droomde van een mooie prins die haar lief had.
Onderweg ontmoette ze een kikker die om een kus vroeg.
Ze vervulde zijn wens en hij werd een knappe prins.

55 De bouwvakkers waren druk bezig met het grote nieuwe huizenblok.
Eerder hadden ze al betonnen geraamtes geplaatst voor de woningen.
Daarna goot eentje cement vanuit een vrachtwagen in de betonmolen.
Vandaag begonnen ze met het metselen van de bakstenen buitenmuur.

56 In Engeland was het koud in de dagen tot kerstmis.
Familie Cooper was bij oma gaan eten op eerste kerstdag.
Iedereen had al cadeautjes gekocht en onder de boom gelegd.
Maar waar ze het meeste naar uitkeken was de kerstpudding.
57 Max had een autosloopbedrijf waar het een grote bende was.
De meeste auto's die hij kreeg waren totaal kapotte wrakken.
Maar op een dag kreeg hij een oude Rolls Royce.
Die verbouwde hij tot de mooiste wagen van het dorp.
58 De barman van de kroeg had weer een drukke avond.
Er waren vele dronken gasten die almaar meer wilden drinken.
Plotseling werd een vent boos en er ontstond een scheldpartij.
Toen hij begon te vechten trapte de barman hem eruit.
59 De Canadese wegen waren helemaal ondergesneeuwd geraakt de afgelopen uren.
Overal waren grote sneeuwschuivers bezig om wegen begaanbaar te houden.
Ook werd met man en macht zout gestrooid tegen gladheid.
Ondanks alles stonden op de snelwegen rond Toronto enorme files.
60 In de oude molen woonde nog altijd een echte molenaar.
De molen werd vrij intensief gebruikt om meel te malen.
Veel gespecialiseerde bakkers gingen gewoonlijk daarheen met hun eigen graan.
Dat werd tussen de stenen fijngemalen tot het gewenste meel.
61 De glazenwasser ging vroeg in de ochtend aan het werk.
Hij schoof zijn ladder uit op straat en klom omhoog.
Fluitend begon hij met het lappen van een grote ruit.
Maar zijn ruitenwisser viel en scheldend klom hij naar beneden.
62 Op zijn vijftigste ging Harry eindelijk weer naar de tandarts.
Hij had zijn leven lang zijn tanden erg slecht verzorgd.
Hij poetste nooit en ging ook niet naar de tandarts.
Helaas kon de tandarts niet anders dan een kunstgebit adviseren.
63 Manuel zag bij de antiquair in zijn straat een schilderijtje.
Het was een mooi kunstwerkje, maar het was erg duur.
Een bevriende kunstkenner schatte het schilderijtje op een lagere prijs.
Na lang onderhandelen kregen ze het voor de lagere prijs.
64 Pompeï was ooit een welvarend Romeins stadje onderaan de Vesuvius.
Maar tijdens een vulkaanuitbarsting raakte het bedolven onder de as.
Al in de achttiende eeuw begonnen archeologen daar met opgravingen.
Tegenwoordig is Pompeï een belangrijke attractie voor toeristen in Italië.
65 Victor was vrachtwagenchauffeur en maakte vele kilometers door heel Europa.
Hij hield van zijn truck en het gevoel van vrijheid.
Bij grote vrachten reed hij altijd met een extra oplegger.
Dat was vaak voordelig vanwege de grotere omzet per rit.
66 In een kolenmijn vond vorige week een grote overstroming plaats.
Op het moment van de overstroming werkten er veertig mijnwerkers.
Een reddingsteam ging de mijn in, op zoek naar overlevenden.
Gelukkig konden alle mijnwerkers levend uit de mijn gehaald worden.
67 In de hoofdstraat van het dorp zat een goede slager.
Het vlees dat hij verkocht kwam van goed verzorgde dieren.
Voor de slacht konden de dieren altijd buiten vrij rondlopen.
Ook kregen ze goed voer, waardoor het vlees beter werd.
68 Salvatore wilde een winkel openen in de Siciliaanse stad Messina.
Maar de maffia maakte hem het leven moeilijk met afpersingen.
Daarom wilde hij getuigen tegen een aangeklaagde maffioso uit Messina.
Maar Salvatore werd vermoord voordat hij zijn getuigenis kon afleggen.
69 Johan was jarenlang imker naast zijn bezigheden in de tuinbouw.
Hij had honingraden staan bij de velden van verschillende boeren.
Jaarlijks haalde hij 's zomers heerlijke verse honing uit de raden.
Ondanks veelvuldige omgang met bijen was hij maar zelden gestoken.

70 In het holst van de nacht was de banketbakker bezig.
Een zakenrelatie van haar had taart besteld voor een bruiloft.
Zijn dochter ging trouwen en zij was dol op gebak.
Daarom liet hij vijf verschillende bruidstaarten aanrukken voor het feest.
71 In de club in het centrum worden grote feesten georganiseerd.
De deejays die er optreden komen van de hele wereld.
Soms komen er later op de avond ook beroemdheden dansen.
De sfeer op de feesten is er echt helemaal geweldig.
72 De trekpaarden van de oude boer hadden nieuwe hoefijzers nodig.
Daarom kwam de hoefsmid afgelopen woensdag langs op zijn boerderij.
Boven vuur smeedde hij hoefijzers op maat voor ieder paard.
Met grote spijkers nagelde hij de ijzers op de hoef.
73 Die student wiskunde werkte op zijn kamer aan zijn scriptie.
Hij probeerde een probleem van de chaostheorie op te lossen.
Wekenlang was hij bezig met het herschrijven van complexe vergelijkingen.
Uiteindelijk lukte het hem toch om zijn stellingen te bewijzen.
74 De nogal onervaren beeldhouwer was een beeld aan het maken.
Uit een blok marmer wilde hij een mooie dame creëren.
Met veel rust en aandacht maakte hij een knap gezicht.
Helaas tikte hij uiteindelijk met zijn hamer de neus eraf.
75 In de herfst ontdekte Kees een vogelnestje in zijn schoorsteen.
Hij liet een schoorsteenveger komen om zijn schoorsteen te legen.
De man haalde er een borstel met een touw doorheen.
Daarna kon Kees weer rustig hout branden in zijn kachel.
76 Enrique ging met vrienden kijken naar een stierengevecht in Spanje.
De enorme stier leek oersterk toen hij de arena binnenkwam.
Maar al gauw werden er spiesen in het beest gestoken.
En zoals gewoonlijk doodde de matador het arme dier tenslotte.
77 Aan de rand van de rivier zaten een paar goudzoekers.
Ze vulden grote schalen met wat modder uit de bedding.
Die draaiden ze in de rondte op zoek naar goud.
Soms vonden ze grotere stukken goud die veel waard zijn.
78 Columbus zocht een nieuwe route om naar Indië te varen.
Met zijn schip voer hij over de hele Atlantische oceaan.
Na een gevaarlijke en wekenlange tocht kwam land in zicht.
Maar anders dan hij vermoedde, had hij nieuw land ontdekt.
79 Ome Keesie hield veel van klussen en bouwde een kast.
Hij zaagde de planken op maat af met een cirkelzaag.
Alle planken pasten mooi en werden goed vastgetimmerd met spijkers.
Maar de laatste plank bleek iets te lang te zijn.
80 Alfred werd verdacht van een gewelddadige overval op een juwelier.
Hij was onschuldig, maar helaas geloofden hem slechts weinig mensen.
Gelukkig had hij een goede advocaat die zeer overtuigend overkwam.
Na een lang en vermoeiend proces werd Alfred uiteindelijk vrijgesproken.

Incoherent stories:

1 Charles verliet zijn vaderland Senegal om in Europa te werken.
Een avond hadden ze een taart achtergelaten in de keuken.
Toevallig kwam een agent de hoek om die hen arresteerde.
Maar na een jaar moest hij al naar de sloop.

2 Een groep hangjongeren maakte zaterdag enorme rotzooi in het parkje.
De wagen zag erg goed uit maar was te duur.
Die nacht sloop een beer binnen die de taart opvrat.
Zijn familie had het geld dat hij stuurde hard nodig.

3 De familie Witteveen verbleef in hun zomerhuis in het woud.
Met een levensgevaarlijk klein bootje werd hij naar Tenerife gesmokkeld.
Twee weken later bleek haar buurman hem te hebben gekocht.
De kwajongens kregen driehonderd uur taakstraf bij de plantsoenendienst opgelegd.

4 Die gladde autoverkoper probeerde Lieneke eens een auto te verkopen.
Twee van hen trapten die avond een bushokje aan diggelen.
Hij moest daar hard werken voor een klein beetje geld.
Pas de volgende ochtend ontdekte de geschrokken vader de inbraak.

5 De reggaeartiest was geboren in een gewelddadige stad op Jamaica.
In het Peugeootje naast hem zag hij een leuke meid.
's Avonds wandelde de oude kromme boer toevallig langs de weide.
Met instemming van de koning werd hij tot ridder geslagen.

6 Net op tijd stopte Jaaps auto voor het rode licht.
Tijdens de veldslag verpletterde Karel vele vijanden met zijn schild.
Dit inspireerde hem bij het schrijven van zijn hoopvolle teksten.
Hij alarmeerde twee dorpsbewoners en kon zo zijn dier redden.

7 De sterke schildknaap Karel diende ridder Archibald met onuitputtelijke kracht.
Het arme beest kon niet meer uit de sloot komen.
Hij reed weg bij groen en zij ging er achteraan.
Zo werd hij op Jamaica populair en brak wereldwijd door.

8 Bij de Rietveldse Weide was een schaap te water geraakt.
Hij had veel agressie meegemaakt en hij blowde uitzonderlijk veel.
Daarom was ridder Archibald zeer tevreden over de dappere jongeman.
Thuis herkende hij haar ineens, ze was zijn nieuwe buurvrouw.

9 Matthias was net uitwonend en vierde dat in een café.
Op vakantie in Zuid-Amerika probeerden ze cocaïne mee te smokkelen.
Hij kon nauwelijks rondkomen en bedelde zijn bestaan bij elkaar.
Toen de jongen werd doodgeschoten hield Mathilde het niet droog.

10 Piet en Jeremy gebruikten veel drugs tijdens hun studie rechten.
Zijn schilderstijl werd nog niet goed begrepen in zijn tijd.
De twee waren lid van verschillende rivaliserende en gewelddadige jeugdbendes.
De politieagent bij het stoplicht zag het door de vingers.

11 Deze bekende Franse schilder trok honderdvijftig jaar geleden door Frankrijk.
De film ging over een onmogelijke verhouding tussen twee geliefden.
Hij reed door rood en z'n lichten waren allebei kapot.
Door hun strafblad nam geen enkel advocatenkantoor hen later aan.

12 Mathilde zat in de bioscoop een romantische film te kijken.
Midden in de nacht fietste hij naar zijn nieuwe huis.
Bij de Nederlandse douane werden ze door een drugshond gesnapt.
Nu betalen rijkelui vele miljoenen euro's voor zijn kleinste werkjes.

13 Al de hele zomer droomde Klaas van schaatsen op natuurijs.
Hij bedreigde de chauffeur en rende weg met zijn geldkistje.
Marloes werd stikjaloers en bazuinde overal rond dat hij vreemdging.
Maar Penelope herkende hem niet meer bij hun eerste ontmoeting.

14 's Avonds stapte een agressieve jongen op de bus richting centrum.
Irene kreeg op haar dertiende een vier jaar oudere vriend.
Na twintig lange jaren rondzwerven kwam Odysseus terug op Ithaka.

Verkleumd kwam hij terug maar was toch voldaan en blij.
15 Irene en Marloes waren sinds de basisschool altijd hartsvriendinnen geweest.
Zijn beeldschone Penelope bleef trouw wachten ondanks de vele huwelijksaanzoeken.
Als eerste schaatste hij op het ijs maar zakte erdoor.
Hij sloeg hem op zijn hoofd en belde de politie.
16 Odysseus verliet zijn jonge vrouw om te vechten tegen Troje.
In januari vroor het voor het eerst sinds jaren hard.
De chauffeur ging achter hem aan met een dikke honkbalknuppel.
Maar hij was echter inderdaad vreemdgegaan en de vriendschap herstelde.
17 Marie-Louise was verlaten door haar partner en stortte helemaal in.
Wereldwijd werd via de media over de rampzalige gevolgen bericht.
's Nachts begonnen zijn honden te blaffen voor een ijsbeer.
Tot haar vreugde had het mengsel een bijzonder lekkere afdronk.
18 In het jaar 1953 vond in Zeeland de watersnoodramp plaats.
Met een hondenslee was hij op zoek gegaan naar voedsel.
Zenuwachtig mengde ze rode en witte wijn in een karaf.
Marie-Louise fleurde weer helemaal op terwijl de therapeut depressief achterbleef.
19 Een Eskimo trok 's winters door het hoge noorden van Groenland.
De gastvrouw schrok toen haar gasten graag rosé wilden drinken.
De therapeut was zelf ook eenzaam en kreeg bijzondere gevoelens.
Zoveel zelfs dat een gedeelte moest worden teruggestuurd of vernietigd.
20 Familie Beerenbrouck kreeg zeer belangrijke gasten te eten op zaterdagavond.
Ze zocht hulp en praatte over haar eenzame droeve leven.
Uit alle windstreken werden heel veel hulpgoederen naar Nederland gestuurd.
Hij doodde het dier en keerde terug naar zijn familie.
21 Johnny werkte in de jaren zestig bij de DAF autofabriek.
Al op zijn twaalfde begon hij met diefstal en moord.
Halverwege zag hij opeens een klein gaatje waaruit water stroomde.
Maar in de Arabische wereld kon hij vrijuit wetenschap bedrijven.
22 Billy the Kid was een jeugdcrimineel uit het Wilde Westen.
Hij was historicus en econoom en grondlegger van de sociologie.
Daarom nam hij stiekem iedere dag een los onderdeel mee.
Hij stak zijn vinger erin en voorkwam zo een dijkbreuk.
23 Khaldun was een belangrijk Arabische geleerde uit de veertiende eeuw.
's Avonds liep hij eens in z'n eentje langs de dijk.
Hij ontsnapte een paar keer uit gevangenissen en werd achtervolgd.
Zo verzamelde hij in vijf jaar een autootje bij elkaar.
24 Hans Brinker was het jongetje dat een belangrijke dijk redde.
Hij droomde van een eigen auto maar had onvoldoende geld.
Wetenschap werd in West-Europa destijds geblokkeerd door de katholieke kerk.
De sheriff doodde hem toen hij nietsvermoedend een saloon binnenstapte.
25 De bankdirecteur wilde een gouden ketting voor zijn vrouw kopen.
Ze pasten met het hele gezin net in de auto.
Maar in zijn kostbare MRI data vond hij goede resultaten.
Hij viel schreeuwend naar beneden en was op slag dood.
26 Walter had weinig tijd om zijn onderzoek af te ronden.
Op een avond probeerde hij via aaneengeknoopte lakens te ontsnappen.
De wagen was tot op het dak volgeladen met bagage.
Helaas moest ze bekennen dat ze de ketting al had.
27 Het gezin ElQasr ging op vakantie naar familie in Marokko.
Hij ging naar een hele dure juwelier in de stad.
Hij klauterde uit het raam maar hoorde een scheurend geluid.
Zo kon hij toch publiceren en beginnen aan nieuw onderzoek.
28 Falco was tot levenslang veroordeeld voor een aantal gruwelijke moorden.
Hij kon die zomer nauwelijks proefpersonen vinden voor zijn MRI-experiment.
Zijn vrouw werd vierentwintig jaar en hij wilde haar verrassen.

Maar na twintig uur rijden ging de auto helaas stuk.
29 De muren van de studeerkamer waren bedekt met dikke boeken.
Drie weken lang was hij van slag door het faillissement.
Ze nodigden hem vaak uit voor een lekker bakje koffie.
Iedereen was er van overtuigd dat het elftal zou winnen.
30 Elke morgen kwam de jonge melkboer vrolijk door de straat.
Toen de beroemde spits het veld betrad juichte het publiek.
Maar hij knapte op en besloot een boerenbedrijf te beginnen.
Zonder aarzelen vonden de handen van de geleerde een boek.
31 Door de kredietcrisis ging de bank waar Jean werkte failliet.
Op de werktafel lagen ook nog torenhoge stapels papieren uitgespreid.
Door de luidsprekers schalden de klanken van het nationale volkslied.
Hij was 's avonds vaak alleen en genoot van hun gezelschap.
32 De elf spelers liepen vol zelfvertrouwen het uitverkochte stadion in.
Alle mensen waar hij melk bracht waren blij met hem.
De professor liep bedachtzaam naar een van de vele boekenkasten.
Na een moeilijke start maakte het bedrijf al snel winst.
33 Nicolaas was bakker in het kleine zaakje op het plein.
Hij kritiseerde de oorlog in TsjetsjeniÃ« en was tegen Poetin.
In de grote bioscopen waren alle voorstellingen bijna helemaal uitverkocht.
De leraar Engels zag dit als bevestiging van zijn onderwijsmethode.
34 Deze populaire Russische journalist schreef kritische artikelen over de regering.
De scholieren waren dus allen zeer benieuwd naar de cijfers.
Daardoor had hij weinig tijd voor z'n vrouw en kinderen.
Ook in de Nijmeegse bioscopen moest men kaartjes vooraf reserveren.
35 Het laatste proefwerk Engels van de brugklas was moeilijk geweest.
De film trok dus heel veel publiek naar de bioscoop.
Ook nam hij geregeld deel aan demonstraties voor meer persvrijheid.
Maar hij bracht altijd vers brood mee voor het ontbijt.
36 De nieuwe film kreeg goede recensies in de nationale pers.
Meestal moest hij 's nachts werken en dan sliep hij overdag.
De cijfers van het tentamen waren gelukkig beter dan verwacht.
Op een ochtend werd hij tijdens zijn dagelijkse wandeling doodgeschoten.
37 Román was een zigeuner en woonde vijf kilometer van Boedapest.
Het zou aan de rand van de grote markt komen.
Sommige van zijn collega's waren daar echter wat makkelijker in.
De eigenaar overwoog het restaurant daarom helemaal niet te openen.
38 Thierry was een heel serieuze wielrenner die meedeed aan de Tour.
De vissoep werd dagelijks vers bereid door de chefkok zelf.
Maar rond de grote markt stonden uitsluitend hele oude gebouwen.
Maar ondanks het geld bleef hij in zijn woonwagen wonen.
39 Het gerenommeerde architectenbureau mocht eindelijk het nieuwe stadhuis gaan ontwerpen.
Hij woonde in een armoedige woonwagen en was heel muzikaal.
Maar vandaag was de chefkok door ziekte helaas niet beschikbaar.
Hij bleef lange tijd fietsen maar won nooit de Tour.
40 Het restaurant aan het kerkplein stond bekend om zijn vissoep.
Hij trainde uitzonderlijk goed en moest niets van doping hebben.
Daardoor werd hij een veel gevraagd violist en verdiende goed.
Het stadhuis moest dus zorgvuldig aan de stijl worden aangepast.
41 Pietje en Klaasje waren stiekem met vuur aan het spelen.
Bijna was het vijf uur en mocht hij naar huis.
Maar zijn stiefmoeder zorgde dat hij geen steentjes meer had.
Zaterdag vertelde ze met veel drama dat ze ging lijnen.
42 Klein Duimpje werd door zijn stiefmoeder naar het bos gestuurd.
Onderweg bleef ze voor de etalage van de banketbakker staan.
Tien minuten later liep hij met zijn koffertje naar huis.

De brandweer moest toen uitrukken om het vuur te blussen.
43 De ambtenaar staarde uit zijn raam op het saaie gemeentehuis.
Ze hadden in de schuur de barbecue met petrolie aangestoken.
Ze twijfelde maar kon de moorkoppen helaas toch niet weerstaan.
Tevergeefs probeerde hij later een spoor van kruimeltjes te maken.
44 Afgelopen vrijdag ging dikke tante Joke zoals altijd boodschappen doen.
Om de weg terug te vinden strooide hij kleine steentjes.
De vlammen werden gauw hoog en waren niet te doven.
Om zes uur at hij thuis zuchtend aardappelen met groente.
45 In de uitgestrekte bossen in Amerika woonde een sterke houthakker.
De wagen was net koel genoeg voor de Nederlandse zomers.
Toen stootte ze per ongeluk een grote stapel blikjes om.
Plotseling viel hij gigantisch uit waarop de meisjes geschrokken wegvluchtten.
46 Gino was een ouderwetse ijscoman met een klein grappig karretje.
Een keurige heer van middelbare leeftijd ergerde zich aan hen.
Met een bevriende houthakker zaagde hij planken van het hout.
Tot haar verassing bleken het de blikjes tomatensoep te zijn.
47 Kassameisjes Mien en Pien zaten te tantebetten in de trein.
Ze zocht blikjes tomatensoep die ze maar niet kon vinden.
Maar op een dag in juli was het uitzonderlijk heet.
Daarvan had hij een blokhut gebouwd waarin hij nu woonde.
48 Een verstrooid omaatje liep met haar winkelwagentje door de supermarkt.
Met een zware bijl hakte hij dagelijks hele bomen om.
De heer kuchte steeds duidelijker maar kreeg geen telkens respons.
Toen besloot hij het smeltende ijs als milkshakes te verkopen.
49 Chagrijnige Harry ging met zijn vrouw op vakantie per auto.
Op droge dagen hooide hij het gras met z'n trekker.
Hij was een erg ijdele maar ook zeer gepassioneerde man.
Hij was vergeten dat de wintertijd was ingegaan die zondag.
50 Een klassiek orkest gaf altijd spetterende optredens in de schouwburg.
Hij had zich verslapen en fietste razendsnel door de stad.
Toen het gras gedroogd was draaide hij er hooibalen van.
Toen zag ze dat zijn kaart vijfentwintig jaar oud was.
51 In de zomer ging boer Piet zijn gras weer maaien.
Hij zat achter het stuur en zijn vrouw las kaart.
Maar bij de tentamenzaal was bij aankomst nog helemaal niemand.
Veel vrouwen uit het orkest hadden een zwak voor hem.
52 Maandagochtend had Michel grote haast omdat hij een tentamen had.
De dirigent van dat orkest was heel bekend in kunstkringen.
In Frankrijk reden ze verkeerd en hij begon te schelden.
Zo had hij wat extra voedsel voor z'n vee 's winters.
53 Vroeger reisde men per trekschuit tussen de steden in Holland.
Ze droomde van een mooie prins die haar lief had.
Daarna goot eentje cement vanuit een vrachtwagen in de betonmolen.
Maar waar ze het meeste naar uitkeken was de kerstpudding.
54 De jonge prinses liep vol verlangen rond de grote vijver.
Eerder hadden ze al betonnen geraamtes geplaatst voor de woningen.
Iedereen had al cadeautjes gekocht en onder de boom gelegd.
Maar het duurde uren voordat je je bestemming bereikt had.
55 De bouwvakkers waren druk bezig met het grote nieuwe huizenblok.
Familie Cooper was bij oma gaan eten op eerste kerstdag.
Je voer over het rustige water van sloten en kanalen.
Ze vervulde zijn wens en hij werd een knappe prins.
56 In Engeland was het koud in de dagen tot kerstmis.
Dit was een soort boot die getrokken werd door paarden.
Onderweg ontmoette ze een kikker die om een kus vroeg.

Vandaag begonnen ze met het metselen van de bakstenen buitenmuur.
57 Max had een autosloopbedrijf waar het een grote bende was.
Er waren vele dronken gasten die almaar meer wilden drinken.
Ook werd met man en macht zout gestrooid tegen gladheid.
Dat werd tussen de stenen fijngemalen tot het gewenste meel.
58 De barman van de kroeg had weer een drukke avond.
Overal waren grote sneeuwschuivers bezig om wegen begaanbaar te houden.
Veel gespecialiseerde bakkers gingen gewoonlijk daarheen met hun eigen graan.
Die verbouwde hij tot de mooiste wagen van het dorp.
59 De Canadese wegen waren helemaal ondergesneeuwd geraakt de afgelopen uren.
De molen werd vrij intensief gebruikt om meel te malen.
Maar op een dag kreeg hij een oude Rolls Royce.
Toen hij begon te vechten trapte de barman hem eruit.
60 In de oude molen woonde nog altijd een echte molenaar.
De meeste auto's die hij kreeg waren totaal kapotte wrakken.
Plotseling werd een vent boos en er ontstond een scheldpartij.
Ondanks alles stonden op de snelwegen rond Toronto enorme files.
61 De glazenwasser ging vroeg in de ochtend aan het werk.
Hij had zijn leven lang zijn tanden erg slecht verzorgd.
Een bevriende kunstkenner schatte het schilderijtje op een lagere prijs.
Tegenwoordig is PompeÃ¯ een belangrijke attractie voor toeristen in ItaliÃ«.
62 Op zijn vijftigste ging Harry eindelijk weer naar de tandarts.
Het was een mooi kunstwerkje, maar het was erg duur.
Al in de achttiende eeuw begonnen archeologen daar met opgravingen.
Maar zijn ruitenwisser viel en scheldend klom hij naar beneden.
63 Manuel zag bij de antiquair in zijn straat een schilderijtje.
Maar tijdens een vulkaanuitbarsting raakte het bedolven onder de as.
Fluitend begon hij met het lappen van een grote ruit.
Helaas kon de tandarts niet anders dan een kunstgebit adviseren.
64 Pompeï was ooit een welvarend Romeins stadje onderaan de Vesuvius.
Hij schoof zijn ladder uit op straat en klom omhoog.
Hij poetste nooit en ging ook niet naar de tandarts.
Na lang onderhandelen kregen ze het voor de lagere prijs.
65 Victor was vrachtwagenchauffeur en maakte vele kilometers door heel Europa.
Op het moment van de overstroming werkten er veertig mijnwerkers.
Voor de slacht konden de dieren altijd buiten vrij rondlopen.
Maar Salvatore werd vermoord voordat hij zijn getuigenis kon afleggen.
66 In een kolenmijn vond vorige week een grote overstroming plaats.
Het vlees dat hij verkocht kwam van goed verzorgde dieren.
Daarom wilde hij getuigen tegen een aangeklaagde maffioso uit Messina.
Dat was vaak voordelig vanwege de grotere omzet per rit.
67 In de hoofdstraat van het dorp zat een goede slager.
Maar de maffia maakte hem het leven moeilijk met afpersingen.
Bij grote vrachten reed hij altijd met een extra oplegger.
Gelukkig konden alle mijnwerkers levend uit de mijn gehaald worden.
68 Salvatore wilde een winkel openen in de Siciliaanse stad Messina.
Hij hield van zijn truck en het gevoel van vrijheid.
Een reddingsteam ging de mijn in, op zoek naar overlevenden.
Ook kregen ze goed voer, waardoor het vlees beter werd.
69 Johan was jarenlang imker naast zijn bezigheden in de tuinbouw.
Een zakenrelatie van haar had taart besteld voor een bruiloft.
Boven vuur smeedde hij hoefijzers op maat voor ieder paard.
De sfeer op de feesten is er echt helemaal geweldig.
70 In het holst van de nacht was de banketbakker bezig.
De deejays die er optreden komen van de hele wereld.
Jaarlijks haalde hij 's zomers heerlijke verse honing uit de raden.

Met grote spijkers nagelde hij de ijzers op de hoef.
71 In de club in het centrum worden grote feesten georganiseerd.
Daarom kwam de hoefsmid afgelopen woensdag langs op zijn boerderij.
Zijn dochter ging trouwen en zij was dol op gebak.
Ondanks veelvuldige omgang met bijen was hij maar zelden gestoken.
72 De trekpaarden van de oude boer hadden nieuwe hoefijzers nodig.
Hij had honingraden staan bij de velden van verschillende boeren.
Soms komen er later op de avond ook beroemdheden dansen.
Daarom liet hij vijf verschillende bruidstaarten aanrukken voor het feest.
73 Die student wiskunde werkte op zijn kamer aan zijn scriptie.
Uit een blok marmer wilde hij een mooie dame creëren.
De man haalde er een borstel met een touw doorheen.
En zoals gewoonlijk doodde de matador het arme dier tenslotte.
74 De nogal onervaren beeldhouwer was een beeld aan het maken.
Hij liet een schoorsteenveger komen om zijn schoorsteen te legen.
Maar al gauw werden er spiesen in het beest gestoken.
Uiteindelijk lukte het hem toch om zijn stellingen te bewijzen.
75 In de herfst ontdekte Kees een vogelnestje in zijn schoorsteen.
De enorme stier leek oersterk toen hij de arena binnenkwam.
Wekenlang was hij bezig met het herschrijven van complexe vergelijkingen.
Helaas tikte hij uiteindelijk met zijn hamer de neus eraf.
76 Enrique ging met vrienden kijken naar een stierengevecht in Spanje.
Hij probeerde een probleem van de chaostheorie op te lossen.
Met veel rust en aandacht maakte hij een knap gezicht.
Daarna kon Kees weer rustig hout branden in zijn kachel.
77 Aan de rand van de rivier zaten een paar goudzoekers.
Hij zaagde de planken op maat af met een cirkelzaag.
Na een gevaarlijke en wekenlange tocht kwam land in zicht.
Na een lang en vermoeiend proces werd Alfred uiteindelijk vrijgesproken.
78 Columbus zocht een nieuwe route om naar Indië te varen.
Hij was onschuldig, maar helaas geloofden hem slechts weinig mensen.
Alle planken pasten mooi en werden goed vastgetimmerd met spijkers.
Soms vonden ze grotere stukken goud die veel waard zijn.
79 Ome Keesie hield veel van klussen en bouwde een kast.
Ze vulden grote schalen met wat modder uit de bedding.
Gelukkig had hij een goede advocaat die zeer overtuigend overkwam.
Maar anders dan hij vermoedde, had hij nieuw land ontdekt.
80 Alfred werd verdacht van een gewelddadige overval op een juwelier.
Met zijn schip voer hij over de hele Atlantische oceaan.
Die draaiden ze in de rondte op zoek naar goud.
Maar de laatste plank bleek iets te lang te zijn.

Appendix B: Chapter 4 Stimuli

Target items:

OORLOG
BIJ
VOGEL
VLINDER
KOE
HOND
EEND
VLO
VOS
KIKKER
PAARD
BOT
WANG
OOR
ELLEBOOG
OOG
GEZICHT
VUIST
HOOFD
RIEM
JAS
KATOEN
JURK
FLES
DOOS
TANTE
BRUID
BROER
DOCHTER
KAARS
VUUR
AARDAPPEL
RIJST
ZOUT
AARDBEI
SUIKER
ZIEKENHUIS
KUSSEN
DAK
KAMER
SCHAAR
ZEEP
LEPEL
HANDDOEK
MUUR
RAAM
RUIMTE
WERELD
KOGEL
POP
VROUW
VERJAARDAG
AARDE
BLOEM
GELD
MOORDENAAR
PAUS
DIEF
BIJBEL
KRANT
LAND
KROON
VLIEGTUIG
FIETS
SNEEUW
ZON
PARAPLU
GOLF
BEZIT
MOGELIJKHEID
VRAAG
REGEL
OPLOSSING
STEM
MODE
GEVAAR
GEZONDHEID
HOEK
EEUW
ZEKERHEID
TIJD
KLEUR
VREDE
SLACHTOFFER
JEUGD
MISBRUIK
KLACHT
KRITIEK
BESCHRIJVING
LEUGEN
TROTS
SCHAAMTE
SLIM
WIJSHEID
ONTDEKKING
ERVARING
GEVOEL
SCHULD
NACHT
ZOMER
HANDTEKENING
LEEUW
AAP
MUIS
VARKEN
KONIJN
HAAI
STAART
SCHILDPAD
WALVIS
VLEUGEL
BEEN
SNOR
MOND
NEUS
HUID
MAAG
DUIM
HOED
SCHOEN
ZIJDE
ROKBROEK
KOOI
VADER
MOEDER
ZOON
TWEELING
OOM
ROOK
WORTEL
KNOFLOOK
CITROEN
VLEES
PERZIK
NAALD
DEKEN
PLAFOND
STOEL
TUIN
HUIS
KEUKEN
MES
SPIEGEL
KANTOOR
STRAND
VELD
BERG
EIGENAAR
FILM
TOUW
JONGEN
MEISJE
TRAAN
MODDER
BOOM
LEGER
SLAGER
KAPITEIN
BOER
VERHAAL
KONING
KONINGIN
AUTO
WOLK
MIST
BLIKSEM
REGEN
WRAAK
VOORDEEL
KUNST
NADEEL
VOORBEELD
TOEKOMST
MEERDERHEID
HITTE
SMAAK
KOORTS
VERSCHIL
AFSTAND
HOOGTE
WETENSCHAP
STILTE
WAARHEID
VIJAND
HELD
BELOFTE
REDEN
TEKEN
STEUN
WENS
SCHOONHEID
EERLIJKHEID
ONSCHULD
GEDULD
GELUK
VERRASSINGGED
ACHTE
TAAL
HERFST

Distractor items:

AANDACHT	VLOED	BETEKENIS	DORP
ONDERWIJS	MAAN	MUZIEK	RICHTING
WET	POTLOOD	DREIGING	PIJL
INGANG	STEEN	VLIEG	KETEN
STIER	DRAAD	SLANG	PISTOOL
AREND	DEEL	NOODZAAK	MEDAILLE
ONZIN	OORZAAK	LICHAAM	LAWAAI
HART	UITSTEL	BORST	VERKOOP
SCHOUDER	GEBEURTENIS	HAAR	SCHANDAAL
GEVANGENIS	DAME	WINKEL	MENIGTE
SJAAL	VOLK	ZAK	REUS
HORLOGE	KEUS	HANDSCHOEN	GUNST
GRAP	MISDAAD	BOODSCHAP	SPRAAK
POST	BESLISSING	LIED	LEEFTIJD
LANDKAART	VREUGDE	ZUS	BOOSHEID
DRUIF	GEHEUGEN	VLAM	ROEM
HONING	MACHT	ERWT	KOMST
SAP	MENING	TAART	KANS
BAD	VOORSTEL	THEE	LES
BUREAU	VUIL	LAKEN	STRO
STOF	WOUD	KRAAN	TABAK
GAZON	POLITIE	AFVAL	RIDDER
VIOOL	SLAAF	WEG	VERPLEEGSTER
HOL	ZEIL	LUCHT	PIRAAT
BOERDERIJ	KAARTJE	STAD	WIEL

Appendix C: Chapter 5 Stimuli

SR/OR/AVR sentences:

Snel verschuilt de tijger, die de jagers bij de bosrand opgemerkt heeft/hebben/hebt, zich tussen de takken van een struik.
In een toespraak roemt de president, die de soldaten tijdens de crisis vertrouwd heeft/hebben/hebt, de steun die hij heeft gekregen.
Achteraf praat de vader, die de zonen bij het concert bewonderd heeft/hebben/hebt, met de dirigent over het optreden.
Vanmorgen had de buurman, die de kwajongens op het plein nageroepen heeft/hebben/hebt, een gesprek met de wijkagent.
Wanhopig schiet de parachutist, die de soldaten bij het gevecht opgemerkt heeft/hebben/hebt, zijn geweer leeg.
Ontroerd luistert de veteraan, die de burgers bij de herdenking gegroet heeft/hebben/hebt, naar het applaus van de aanwezigen.
Om vijf uur vertrekt de koningin, die de burgemeesters op de receptie begroet heeft/hebben/hebt, naar het paleis in Den Haag.
Tijdens het vertrek staat de kapitein, die de matrozen bij het laden geholpen heeft/hebben/hebt, op de brug van het schip.
Vanwege de nederlaag vreest de trainer, die de spelers na de wedstrijd afgekraakt heeft/hebben/hebt, dat zijn ontslag nu wel snel zal komen.
Na de lessen heeft de leraar, die de leerlingen tijdens de les gekwetst heeft/hebben/hebt, een gesprek met de directeur.
Kwaad leest de huisbaas, die de huurders bij de betaling misleid heeft/hebben/hebt, de oproep voor de rechtszaak.
Na het festival wil de uitgever, die de dichters over het optreden geschreven heeft/hebben/hebt, een serie dichtbundels gaan uitgeven.
Gespannen wacht de patiënt, die de verpleegsters op de afdeling gesproken heeft/hebben/hebt, op de uitslag van het onderzoek.
Uitvoerig vertelt de woordvoerder, die de journalisten na de crisis opgebeld heeft/hebben/hebt, hoe de zaken ervoor staan.
In een brief doet de zwemster, die de juryleden na de wedstrijd gemeden heeft/hebben/hebt, haar beklag bij de bond.
Met veel plezier denkt de oma, die de kleinkinderen op het feest toegelachen heeft/hebben/hebt, aan het feest terug.
In haar rede prijst de koningin, die de politici op het feest toegesproken heeft/hebben/hebt, de architectuur van het gebouw.
Woedend loopt de directeur, die de stakers bij de poort aangesproken heeft/hebben/hebt, naar zijn auto terug.
Na de rechtszaak staat de aanklager, die de criminelen tijdens de zitting aangehoord heeft/hebben/hebt, de pers te woord.
In de vergadering legt de dwarsligger, die de bestuursleden in een brief bekritiseerd heeft/hebben/hebt, de actie uit aan de leden.
Op het politiebureau leggen de agenten, die de demonstrant bij de betoging geslagen heeft/hebben/hebt, een verklaring af tegenover de rechercheur.
Ongerust kijken de hardlopers, die de wandelaar in het park gegroet heeft/hebben/hebt, naar de regenwolken in de lucht.
Bij de buluitreiking zijn de professoren, die de student na de studie opgehemeld heeft/hebben/hebt, vol lof over de scriptie.
Altijd hebben de heren, die de knecht bij het werk gesteund heeft/hebben/hebt, veel plezier in hun bezigheden.
Goedgehumeurd verlaten de chefs, die de medewerker op de receptie gefeliciteerd heeft/hebben/hebt, de kantine van het bedrijf.
Na een tijdje gaan de baby's, die de ouder bij het verschonen aangekeken heeft/hebben/hebt, weer lekker slapen in hun wiegjes.
Thuis nemen de grimeurs, die de toneelspeler na het stuk gecomplimenteerd heeft/hebben/hebt, een borrel voor het slapen gaan.
Uitvoerig bespreken de ministers, die de ambtenaar op de vergadering verwelkomd heeft/hebben/hebt, de taken die er liggen.
Tegen hun superieuren durven de agenten, die de spion in de hoofdstad geschaduwd heeft/hebben/hebt, niet te bekennen dat zij fouten hebben gemaakt.
Vlak na middernacht gaan de kelners, die de gast na de maaltijd gegroet heeft/hebben/hebt, nog wat drinken in een café.
Na lang aarzelen vertellen de misdadigers, die de advocaat over het proces gesproken heeft/hebben/hebt de toedracht van de moord.
Binnenkort krijgen de kunstenaars, die de vrouw tijdens het bezoek bewonderd heeft/hebben/hebt, een atelier in het centrum.
Na de voorstelling staan de clowns, die het kind tijdens de act uitgelachen heeft/hebben/hebt, bij de uitgang van de circustent.
Snel duiken de cowboys, die de indiaan op de steppe beschoten heeft/hebben/hebt, achter een rotsblok om zich tegen de pijlen te beschermen.
Volgens de berichten verlaten de eigenaars, die de pachter over de betaling gebeld heeft/hebben/hebt, binnenkort het landgoed.
Op de televisie zien de supporters, die de scheidsrechter op de radio bekritiseerd heeft/hebben/hebt, de beelden van de wedstrijd.
Bij de evaluatie hebben de bedrijfsleiders, die de stagiair bij de stage geholpen heeft/hebben/hebt, veel kritiek op de resultaten.
Bezorgd bekijken de examinatoren, die de cursist over het tentamen gesproken heeft/hebben/hebt, de cijfers die behaald zijn.
Geduldig helpen de verkoopsters, die de klant op de kledingafdeling gezien heeft/hebben/hebt, bij de keuze van een jas.
Vanwege het onderzoek moeten de inbrekers, die de bewoner bij de inbraak neergeslagen heeft/hebben/hebt, nog een tijdje op het politiebureau blijven.
Woedend leest de dansleraar, die de dansers bij het optreden bewonderd heeft/hebben/hebt, de recensie in de krant.
Gehaast geeft de commissaris, die de rechercheurs op het bureau gezocht heeft/hebben/hebt, de opdracht aan zijn mannen.
Opgewekt vertelt de gastvrouw, die de vriendinnen bij de voordeur omhelsd heeft/hebben/hebt, over de reis die ze heeft gemaakt.
Onwillig geeft de wethouder, die de raadsleden in het debat gehekeld heeft/hebben/hebt, openheid van zaken over de fraude.
Na de ontmoeting gaat de paus, die de chirurgen in het ziekenhuis gesproken heeft/hebben/hebt, naar zijn woning in het Vaticaan.
Kwajongensachtig bedreigde de admiraal, die de generaals tijdens de slag gered heeft/hebben/hebt, de maarschalk met zijn zwaard.
Ontzet keurde de stylist, die de visagistes voor de bruiloft geadviseerd heeft/hebben/hebt, het werk van de kapster.
Opgewekt schopte de spits, die de middenvelders tijdens de wedstrijd afgeblaft heeft/hebben/hebt, de bal recht naar de keeper.
Hartelijk bedankte de Japanner, die de toeristen voor een kunstwerk gefotografeerd heeft/hebben/hebt t, de Amerikaan voor zijn hulp.
Nonchalant kocht de premier, die de chauffeurs na de vergadering gehekeld heeft/hebben/hebt, de maffiabaas om.

Zenuwachtig liet de zanger, die de pianisten op het podium begeleid heeft/hebben/hebt, de harpist aan zijn lot over.
Sceptisch hielp de conductrice, die de machinisten naar de spoorbaan geroepen heeft/hebben/hebt, de spoorwegbeambte met de uitleg van de nieuwe route.
Ongemerkt beconcurreerde de waarzegster, die de handlezers op de markt opgehemeld heeft/hebben/hebt, de helderziende uit het buurdorp.
Helaas bedroog de moeder, die de dochters aan de deur begroet heeft/hebben/hebt, de nicht van haar vriendin.
Alweer kleineerde de kok, die de afwassers in de keuken geërgerd heeft/hebben/hebt, het hulpje vanwege zijn werk.
Hevig bedreigde de Hagenees, die de Amsterdammers op het plein bespuugd heeft/hebben/hebt, de Rotterdammers tijdens de wedstrijd.
Fel beledigde de Utrechter, die de Groningers tijdens het gesprek afgekeurd heeft/hebben/hebt, de Leidenaar die de Groningers verdedigde.
Meedogenloos kraakte de bankdirecteur, die de zakenmannen door een schuldeis geruïneerd heeft/hebben/hebt, de aandeelhouder van het bedrijf af.
Meelevend troostte de eenhoorn, die de kabouters in het bos bekeken heeft/hebben/hebt, de fee die helemaal wanhopig was.
Ruig moedigde de rokkenjager, die de macho's met wilde verhalen geinspireerd heeft/hebben/hebt, op het feest de vrijgezel aan.
Om zeven uur gaan de rechercheurs, die de verdachte bij het verhoor beledigd heeft/hebben/hebt, naar huis om te eten.
Vrolijk laten de conducteurs, die de passagier bij het instappen toegelachen heeft/hebben/hebt, de trein vertrekken.
Opgeruimd ontvangen de organisatoren, die de artiest in het verleden gekend heeft/hebben/hebt, de gasten bij de afscheidsreceptie in het hotel.
Ongerust ontvangen de artsen, die de accountant over de afrekening geschreven heeft/hebben/hebt, de advocaat in de vergaderkamer op de bovenste etage.
Uiteindelijk ontvangen de detectives, die de drugsbaron over de concurrent gesproken heeft/hebben/hebt, de getuige op een zeer geheime plek.
Onverstoorbaar negeerden de oude mannen, die de vrouw uit de zee gered heeft/hebben/hebt, de brieven van het jonge ding.
Dwangmatig pestten de bakkers, die de slager voor de markt genegeerd heeft/hebben/hebt, de groenteboer uit zijn winkel.
Overtuigend ondersteunden de brandweermannnen, die de agent na de brand geroepen heeft/hebben/hebt, de uitleg van de verpleger.
Ontevreden betaalden de bouwvakkers, die de voorman met de werktijd belazerd heeft/hebben/hebt, de aannemer het verschuldigde bedrag.
Geduldig troostten de chauffeurs, die de politieagent in het kantoor bedankt heeft/hebben/hebt, de lifter over het ongeluk met zijn vriendin.
Na veel frustratie mopperden de schilders, die de huisbaas met het project geholpen heeft/hebben/hebt, tenslotte op de huurder.
Ondanks alle problemen negeerden de assistenten, die de hoogleraar over het experiment geschreven heeft/hebben/hebt, de waarschuwingen van de student.
Spoedig haalden de olifanten, die de neushoorn op de vlakte gepasseerd heeft/hebben/hebt, de zebra in.
Stiekem deden de priesters, die de predikant na de mis omhelsd heeft/hebben/hebt, de kapelaan van de parochie na.
Snel kropen de oorwurmen, die de spin in de tuin beklommen heeft/hebben/hebt, over de duizendpoot het gras in.
Heel snel reanimeerden de broers, die de zuster na het ongeluk geholpen heeft/hebben/hebt, het zwaar gewonde nichtje.
Gewetenloos logen de hertogen, die de baron tijdens het banket afgescheept heeft/hebben/hebt, tegen de graaf uit het buurland.
Vrolijk bespotten de cabaretiers, die de poppenspeler tijdens de toneelavond genegeerd heeft/hebben/hebt, het publiek met gemene grappen.
Bewust negeerden de machinisten, die de conducteur op het perron begroet heeft/hebben/hebt, de reiziger met de zware koffer.
In de ochtend belden de loodgieters, die de melkboer met de offerte misleid heeft/hebben/hebt, de klusjesman op zijn kantoor.
Tijdens het debat ergerde de politicus, die de diplomaten in een toespraak geantwoord heeft/hebben/hebt, de minister van Buitenlandse Zaken.
Uitvoerig informeerde de Australiër, die de Aziaten vanuit de hoofdstad gebeld heeft/hebben/hebt, de Europeaan over de zaak.
Heel verrassend deed de sopraan, die de alten van het koor bekeken heeft/hebben/hebt, de tenor in elk detail na.
Na de conferentie minachtte de natuurkundige, die de biologen van de universiteit geschreven heeft/hebben/hebt, de scheikundige die bij een bedrijf werkt.
Heel verrassend liet de psycholoog, die de sociologen bij het onderzoek geholpen heeft/hebben/hebt, de pedagoog in de discussie winnen.
Onverwachts overlegde de voogd, die de leraren met het rapport overvallen heeft/hebben/hebt, met de leerlingen van de klas.
Volgens de krant aanbad de fotograaf, die de modellen in de kleedkamer verleid heeft/hebben/hebt, de designer van de nieuwe collectie.
Listig lonkte de socialist, die de liberalen tijdens het debat geplaagd heeft/hebben/hebt, onverwachts naar de democraat.
Bruusk overviel de romanticus, die de realisten tijdens de expositie uitgescholden heeft/hebben/hebt, het atelier van de impressionist.
Verbitterd schold het omaatje, die de zusters uit het klooster vermaakt heeft/hebben/hebt, vanwege de erfenis op de familie.
Bars onderwees de cateraar, die de klanten in de eetzaal gecommandeerd heeft/hebben/hebt, voor de derde maal de stagiair.
Dolgelukkig bedankte de bruid, die de bruidsmeisjes na de ceremonie gekust heeft/hebben/hebt, de bruidegom voor de ring.
Jammer genoeg beet de labrador, die de herders in het veld besmet heeft/hebben/hebt, in de benen van het schaap.
Grootmoedig hielp de kapitein, die de piraten op volle zee verslagen heeft/hebben/hebt, het schip van de zeerover.

Resoluut negeerde de kaper, die de terroristen op klaarlichte dag verraden heeft/hebben/hebt, het voorstel van de onderhandelaar.
Opgewonden verliet de stewardess, die de piloten in de kleedkamer bekeken heeft/hebben/hebt, de kamer van de steward.
Roekeloos negeerde de patiënt, die de verzorgers na een uur gewaarschuwd heeft/hebben/hebt, de oproep van de bewaking.
Onaangedaan bevrijdde de commando, die de guerrilla's met een hinderlaag gevangen heeft/hebben/hebt, in de jungle de gevangene.
Achter de schermen beïnvloedt de commissie, die de studenten in de rechtszaal bedrogen heeft/hebben/hebt, het oordeel van de adviseur.
Liefdevol accepteert de romanticus, die de geliefden voor het eerst geschilderd heeft/hebben/hebt, een gift van de minnares.
Ongerust ontweken de zwervers, die de alcoholist op het plein bestolen heeft/hebben/hebt, de junkie die geen geld had.
Natuurljk waardeerden de schaatsers, die de tennisser bij de prijsuitreiking vertrouwd heeft/hebben/hebt, ook de zeer bekende golfer.
Brutaal snauwden de lassers, die de metselaar op de bouwplaats gecommandeerd heeft/hebben/hebt, de schilder af die een fout had gemaakt.
Gelukkig hielpen de violisten, die de saxofonist tijdens de repetitie aangemoedigd heeft/hebben/hebt, de drummer van de band.
Tijdens de voorstelling bestudeerden de clowns, die de goochelaar van het circus geamuseerd heeft/hebben/hebt, de trucs van de acrobaat.
Tijdens de repetitie beledigden de figuranten, die de regisseur van het stuk ontmoedigd heeft/hebben/hebt, de heel beroemde hoofdrolspeler.
Spottend vermaakten de cipiers, die de gevangenen in de recreatieruimte voorgelezen heeft/hebben/hebt, de getuige met hun sleutelbos.
Opgelucht hielpen de piloten, die de instructeur na het examen bedankt heeft/hebben/hebt, de vrouw van de ingenieur.
Ontgoocheld ontsloegen de atleten, die de jogger met het trainingsschema geholpen heeft/hebben/hebt, de onkundige therapeut.
Na lang beraad namen de ingenieurs, die de jurist op staande voet ontslagen heeft/hebben/hebt, de steekpenningen van de fraudeurs aan.
Wellustig versierden de gokkers, die het blondje in de slaapkamer ontdekt heeft/hebben/hebt, de vrouw van de animator.
Op intuitie onderhandelden de antiquairs, die de verzamelaar in de haast vergeten heeft/hebben/hebt, over grote bedragen met de taxateurs.
Geërgerd duwen de portiers, die de hooligan in de nachtclub geweigerd heeft/hebben/hebt, de bezoeker naar achteren.
Glashard negeerden de prinsessen, die de held uit de sloppenwijk bedankt heeft/hebben/hebt, de blikken van de prins.
Bikkelhard schoppen de verzorgers, die de bokser uit de ring geslagen heeft/hebben/hebt, in het kruis van de scheidsrechter.
Gelukkig ontlopen de sporters, die de verslaggever van de krant geschopt heeft/hebben/hebt, de verwensingen van de fan.
Noodgedwongen vermaken de dwergen, die de kabouter uit de grot gedragen heeft/hebben/hebt, het zeerlieve aardmannetje.
Luid brullend vernederen de bullebakken, die de grapjas in zijn ziel gekwetst heeft/hebben/hebt, het personeel van de kroegbaas.
Zonder scrupules plezieren de serveersters, die de klant in de nachtclub versierd heeft/hebben/hebt, vervolgens de manager.
Morrend accepteren de bankemployees, die de cliënt in de kluis geobserveerd heeft/hebben/hebt, het gezag van de leider.

References

Alvarez, P., & Squire, L. R. (1994). Memory consolidation and the medial temporal lobe: a simple network model. *Proceedings of the National Academy of Sciences*, *91*(15), 7041-7045.

Angrilli, A., Penolazzi, B., Vespignani, F., De Vincenzi, M., Job, R., Ciccarelli, L., Palomba, D., & Stegagno, L. (2002). Cortical brain responses to semantic incongruity and syntactic violation in Italian language: an event-related potential study. *Neuroscience Letters*, *322*(1),

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (Release 2) [CD-ROM]. Philidelphia: Linguistic Data Consortium, University of Pennsylvania.

Barber, H., & Carreiras, M. (2005). Grammatical Gender and Number Agreement in Spanish: An ERP Comparison. *Journal of Cognitive Neuroscience*, *17*(1), 137–153.

Bastiaansen, M. C. M., & Knösche, T. R. (2000). Tangential derivative mapping of axial MEG applied to event-related desynchronization research. *Clinical Neurophysiology*, *111*(7), 1300–1305.

Bastiaansen, M. C. M., & Hagoort, P. (2006). Oscillatory neuronal dynamics during language comprehension. *Progress in Brain Research*, *159*(06), 179–196.

Bastiaansen, M. C. M., & Hagoort, P. (2015). Frequency-based Segregation of Syntactic and Semantic Unification during Online Sentence Level Language Comprehension. *Journal of Cognitive Neuroscience*, *27*(11), 2095–2107.

Bastiaansen, M. C. M., Magyari, L., & Hagoort, P. (2010). Syntactic unification operations are reflected in oscillatory dynamics during on-line sentence comprehension. *Journal of Cognitive Neuroscience*, *22*(7), 1333–1347.

Bastiaansen, M. C. M., Mazaheri, A., & Jensen, O. (2012). Beyond ERPs: oscillatory neuronal dynamics. In S. Luck & E. Kappenman (Eds.), *Oxford handbook of event-related potential components* (pp. 31-49). New York: Oxford University Press.

Bastiaansen, M. C. M., Oostenveld, R., Jensen, O., & Hagoort, P. (2008). I see what you mean: theta power increases are involved in the retrieval of lexical semantic information. *Brain and Language*, *106*(1), 15–28.

Bastiaansen, M. C. M., van Berkum, J. J. A, & Hagoort, P. (2002). Event-related theta power increases in the human EEG during online sentence processing. *Neuroscience Letters*, *323*(1), 13–16.

Bastiaansen, M. C. M., van Berkum, J. J. A., & Hagoort, P. (2002). Syntactic Processing

Modulates the θ Rhythm of the Human EEG. *NeuroImage*, *17*(3), 1479–1492.

Bastiaansen, M. C. M., van der Linden, M., ter Keurs, M., Dijkstra, T., & Hagoort, P. (2005). Theta Responses Are Involved in Lexical-Semantic Retrieval during Language Processing. *Journal of Cognitive Neuroscience*, *17*(3), 530–541.

Bastos, A. M., Usrey, W. M., Adams, R. a, Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, *76*(4), 695–711.

Betancort, M., Carreiras, M., & Sturt, P. (2009). The processing of subject and object relative clauses in Spanish: An eye-tracking study. *The Quarterly Journal of Experimental Psychology*, *62*(10), 1915–1929.

Blair, R. C., & Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology*, *30*(5), 518-524.

Brennan, J., Lignos, C., Embick, D., & Roberts, T. P. L. (2014). Spectro-temporal correlates of lexical access during auditory lexical decision. *Brain and Language*, *133*, 39–46.

Bressler, S. L., & Richter, C. G. (2014). Interareal oscillatory synchronization in top-down neocortical processing. *Current Opinion in Neurobiology*, *31C*, 62–66.

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: rethinking the functional role of the P600 in language comprehension. *Brain Research*, *1446*, 127–143.

Bubic, A., von Cramon, D. Y., & Schubotz, R. I. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience*, *4*, 25.

Caplan, D., Alpert, N., & Waters, G. (1998). Effects of Syntactic Structure and Propositional Number on Patterns of Regional Cerebral Blood Flow. *Journal of Cognitive Neuroscience*, *10*(4), 541–552.

Caramazza, A., & Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language*, *3*(4), 572–582.

Carreiras, M., Duñabeitia, J. A., Vergara, M., de la Cruz-Pavía, I., & Laka, I. (2010). Subject relative clauses are not universally easier to process: Evidence from Basque. *Cognition*, *115*(1), 79–92.

Catani, M., Mesulam, M. M., Jakobsen, E., Malik, F., Martersteck, A., Wieneke, C., … Rogalski, E. (2013). A novel frontal pathway underlies verbal fluency in primary progressive aphasia. *Brain : A Journal of Neurology*, *136*(Pt 8), 2619–28.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MT: MIT Press.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*(01), 3–42.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *The Behavioral and Brain Sciences*, *36*(3), 181-204.

Cohen, L., & Mehler, J. (1996). Click monitoring revisited: An on-line study of sentence comprehension. *Memory & Cognition*, *24*(1), 94–102.

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the Unexpected: Event-related Brain Response to Morphosyntactic Violations. *Language and Cognitive Processes*, *13*(1), 21–58.

Davidson, D. J., & Indefrey, P. (2007). An inverse relation between event-related and time-frequency violation responses in sentence processing. *Brain Research*, *1158*, 81–92.

Dell, G. S., & Chang, F. (2014). The P-chain: relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1634), 20120394.

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

Dewaele, J., & Véronique, D. (2001). Gender assignment and gender agreement in advanced French interlanguage: a cross-sectional study. *Bilingualism: Language and Cognition*, *4*(03), 275–297.

Dowens, M. G., Guo, T., Guo, J., Barber, H., & Carreiras, M. (2011). Gender and number processing in Chinese learners of Spanish – Evidence from Event Related Potentials. *Neuropsychologia*, *49*(7), 1651–1659.

Dowens, M. G., Vergara, M., Barber, H. A., & Carreiras, M. (2010). Morphosyntactic Processing in Late Second-Language Learners. *Journal of Cognitive Neuroscience*, *22*(8), 1870–1887.

Elman, J. L. (1991). Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning*, *7*(2/3), 195–225.

Engel, A. K., & Fries, P. (2010). Beta-band oscillations--signalling the status quo? *Current Opinion in Neurobiology*, *20*(2), 156–165.

Federmeier, K. D. (2007). Thinking ahead: the role and roots of prediction in language comprehension. *Psychophysiology*, *44*(4), 491-505.

Fedorenko, E., & Thompson-Schill, S. L. (2014). Reworking the language network. *Trends in Cognitive Sciences*, *18*(3), 120–126.

Ferreira, F., & Patson, N. D. (2007). The 'good enough' approach to language comprehension. *Language and Linguistics Compass*, *1*(1-2), 71–83.

Ford, M. (1983). A method for obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior*, *22*(2), 203–218.

Foucart, A., & Frenck-Mestre, C. (2012). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1–L2 syntactic similarity. *Bilingualism: Language and Cognition*, *15*(1), 202.

Fox, B. A., & Thompson, S. A. (1990). A Discourse Explanation of the Grammar of Relative Clauses in English Conversation. *Language*, *66*(2), 297.

Frauenfelder, U., Segui, J., & Mehler, J. (1980). Monitoring around the relative clause. *Journal of Verbal Learning and Verbal Behavior*, *19*(3), 328–337.

Frazier, L. (1987). Syntactic processing: Evidence from dutch. *Natural Language and Linguistic Theory*, *5*(4), 519–559.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive psychology, 14*(2), 178-210.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2), 78–84.

Friederici, A. D. (2009). Pathways to language: fiber tracts in the human brain. *Trends in Cognitive Sciences*, *13*(4), 175–181.

Friederici, A. D., & Meyer, M. (2004). The brain knows the difference: two types of grammatical violations. *Brain Research*, *1000*(1-2), 72–77.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, *1*(3), 183–192.

Friederici, A. D., & Singer, W. (2015). Grounding language processing on basic neurophysiological principles. *Trends in Cognitive Sciences*, *19*(6), 329–338.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1456), 815–836.

Friston, K. J., Bastos, A. M., Pinotsis, D., & Litvak, V. (2014). LFP and oscillations-what do they tell us? *Current Opinion in Neurobiology*, *31C*, 1–6.

Gennari, S. P., & MacDonald, M. C. (2008). Semantic indeterminacy in object relative clauses. *Journal of Memory and Language*, *58*(2), 161–187.

Gennari, S. P., & MacDonald, M. C. (2009). Linking production and comprehension processes: The case of relative clauses. *Cognition*, *111*(1), 1–23.

Gevins, A. (1997). High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, *7*(4), 374–385.

Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition*, *68*(1), 1–76.

Gibson, E. (2000). The Dependency Locality Theory : A Distance -Based Theory of Linguistic Complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain* (pp. 95–126). Cambridge, MA: MIT Press.

Gibson, E., & Wu, H.-H. I. (2013). Processing Chinese relative clauses in context. *Language and Cognitive Processes*, *28*(1-2), 125–155.

Giraud, A., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517.

Glenberg, A. M., Becker, R., Klötzer, S., Kolanko, L., Müller, S., & Rinck, M. (2009). Episodic affordances contribute to language comprehension. *Language and Cognition*, *1*(1), 113-135.

Goldberg, A. E. (2003). Constructions: a new theoretical approach to language. *Trends in Cognitive Sciences*, *7*(5), 219–224.

Gordon, P. C., & Hendrick, R. (2004). Relativization, ergativity, and corpus frequency. *Linguistic Inquiry*, (1987), 456–463.

Gordon, P. C., & Lowder, M. W. (2012). Complex Sentence Processing: A Review of Theoretical Perspectives on the Comprehension of Relative Clauses. *Language and Linguistics Compass*, *6*(7), 403–415.

Gouvea, A. C., Phillips, C., Kazanina, N., & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes*, *25*(2), 149–188.

Gratton, G., Coles, M. G. ., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 468–484.

Gray, C., König, P., Engel, A., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Letters to Nature*, *338*, 334–337.

Grodner, D., & Gibson, E. (2005). Consequences of the Serial Nature of Linguistic Input for Sentenial Complexity. *Cognitive Science*, *29*(2), 261–290.

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: a critical tutorial review. *Psychophysiology*, *48*(12), 1711-1725.

Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic Gender and Semantic Expectancy: ERPs Reveal Early Autonomy and Late Interaction. *Journal of Cognitive Neuroscience*, *12*(4), 556–568.

Gunter, T. C., Stowe, L. A., & Mulder, G. (1997). When syntax meets semantics. *Psychophysiology*, *34*(6), 660–676.

Guillelmon, D., & Grosjean, F. (2001). The gender marking effect in spoken word recognition: The case of bilinguals. *Memory & Cognition*, *29*(3), 503–511.

Gross, J., Kujala, J., Hamalainen, M., Timmermann, L., Schnitzler, A., & Salmelin, R. (2001). Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(2), 694–699.

Haarmann, H. I., Cameron, K. a, & Ruchkin, D. S. (2002). Neural synchronization mediates on-line sentence processing: EEG coherence evidence from filler-gap constructions. *Psychophysiology*, *39*(6), 820–5.

Hagoort, P. (2003). How the brain solves the binding problem for language: a neurocomputational model of syntactic processing. *NeuroImage*, *20*, S18–S29.

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in Cognitive Sciences*, *9*(9), 416–423.

Hagoort, P. (2013). MUC (Memory, Unification, Control) and beyond. *Frontiers in Psychology*, *4*, 416.

Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca's region and beyond. *Current Opinion in Neurobiology*, *28*, 136–141.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an erp measure of syntactic processing. *Language and Cognitive Processes*, *8*(4), 439–483.

Hagoort, P., & van Berkum, J. (2007). Beyond the sentence given. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1481), 801–811.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, *11*(2), 194–205.

Hald, L. A., Bastiaansen, M. C. M., & Hagoort, P. (2006). EEG theta and gamma responses to semantic violations in online sentence processing. *Brain and Language*, *96*(1), 90–105.

Hale, J. (2001). A probabilistic earley parser as a psycholinguistic model. In *Second meeting of the North American Chapter of the Association for Computational Linguistics on Language technologies 2001 - NAACL '01* (Vol. 2, pp. 1–8). Morristown, NJ, USA: Association for Computational Linguistics.

Hanslmayr, S., Spitzer, B., & Bäuml, K. (2009). Brain oscillations dissociate between semantic and nonsemantic encoding of episodic memories. *Cerebral Cortex*, *19*(7), 1631-1640.

Hanslmayr, S., Staudigl, T., & Fellner, M.-C. (2012). Oscillatory power decreases and long-term memory: the information via desynchronization hypothesis. *Frontiers in Human Neuroscience*, *6*(74), 1–12.

Herrmann, C. S. (2001). Human EEG responses to 1-100 Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Experimental Brain Research*, *137*(3-4), 346-353.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393–402.

Hipp, J. F., & Siegel, M. (2013). Dissociating neuronal gamma-band activity from cranial and ocular muscle activity in EEG. *Frontiers in Human Neuroscience*, *7*, 338.

Hoffmann, C (2011). Do neuronal oscillations reflect semantic coherence in short discourses? Unpublished master's thesis, Radboud University, Nijmegen, (Supervisors: Bastiaansen, M. C. M. & Schriefers, H.).

Holle, H., Obermeier, C., Schmidt-Kassow, M., Friederici, A. D., Ward, J., & Gunter, T. C. (2012). Gesture Facilitates the Syntactic Analysis of Speech. *Frontiers in Psychology*, *3*, 1–12.

Holmes, V. M., & Dejean de la Bâtie, B. (1999). Assignment of grammatical gender by native speakers and foreign learners of French. *Applied Psycholinguistics*, *20*(04), 479–506.

Holmes, V. M., & O'Regan, J. K. (1981). Eye fixation patterns during the reading of relative-clause sentences. *Journal of Verbal Learning and Verbal Behavior*, *20*(4), 417–430.

Hsiao, F., & Gibson, E. (2003). Processing relative clauses in Chinese. *Cognition*, *90*(1), 3–27.

Huettig, F. (2015). Four central questions about prediction in language processing. *brain research*, *1626*, 118-135.

Huettig, F., & Mani, N. (2016). Is prediction necessary to understand language? Probably not. *Language, Cognition and Neuroscience*, *31*(1), 19-31.

Jackendoff, R. (2007). A Parallel Architecture perspective on language processing. *Brain Research*, *1146*, 2–22.

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in Human Neuroscience*, *4*, 186.

Johnson, J. D., McDuff, S. G. R., Rugg, M. D., & Norman, K. A. (2009). Recollection, Familiarity, and Cortical Reinstatement: A Multivoxel Pattern Analysis. *Neuron*, *63*(5), 697-708.

Johnson, J. D., & Rugg, M. D. (2007). Recollection and the Reinstatement of Encoding-Related Cortical Activity. *Cerebral Cortex*, *17*(11), 2507-2515.

Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2000). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clinical Neurophysiology*, *111*(10), 1745–1758.

Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. Trends in Cognitive Sciences, 9(11), 512–518.

Just, M.A., Carpenter, P. A., Keller, T. A., Eddy, W. F., & Thulborn, K. R. (1996). Brain Activation Modulated by Sentence Comprehension. *Science*, *274*(5284), 114–116.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122–149.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*(2), 159–201.

Kahana, M. J., Seelig, D., & Madsen, J. R. (2001). Theta returns. *Current Opinion in Neurobiology*, *11*(6), 739–744.

Kielar, A., Meltzer, J. A., Moreno, S., Alain, C., & Bialystok, E. (2014). Oscillatory Responses to Semantic and Syntactic Violations. *Journal of Cognitive Neuroscience*, *26*(12), 2840–

2862.

Kielar, A., Panamsky, L., Links, K. A., & Meltzer, J. A. (2015). Localization of electrophysiological responses to semantic and syntactic anomalies in language comprehension with MEG. *NeuroImage*, *105*, 507–524.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, *30*(5), 580–602.

King, J. W., & Kutas, M. (1995). Who Did What and When? Using Word- and Clause-Level ERPs to Monitor Working Memory Usage in Reading. *Journal of Cognitive Neuroscience*, *7*(3), 376–395.

König, P., & Schillen, T. (1991). Stimulus-dependent assembly formation of oscillatory responses: I. Synchronization. *Neural Computation*, *3*, 155–166.

Koornneef, A. W., & Sanders, T. J. M. (2013). Establishing coherence relations in discourse: The influence of implicit causality and connectives on pronoun resolution. *Language and Cognitive Processes*, *28*(8), 1169–1206.

Kopell, N., Whittington, M. A., & Kramer, M. A. (2011). Neuronal assembly dynamics in the beta1 frequency range permits short-term memory. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(9), 3779–84.

Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain and Language*, *109*(2-3), 68–74.

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension?. *Language, cognition and neuroscience*, *31*(1), 32-59.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annual Review of Psychology, 62, 621–647.

Kutas, M., & Hillyard, S. (1984). Brain potentials during reading reflect word expectancy and semantic association. Nature, 307, 161–163.

Kveraga, K., Ghuman, A. S., & Bar, M. (2007). Top-down predictions in the cognitive brain. *Brain and Cognition*, *65*(2), 145-168.

Kwon, N., Gordon, P. C., Lee, Y., Kluender, R., & Polinsky, M. (2010). Cognitive and linguistic factors affecting subject/object asymmetry: An eye-tracking study of prenominal relative clauses in Korean. *Language*, *86*(3), 546–582.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. Discourse Processes, 25(2-3), 259–284.Lau, E., Stroud, C., Plesch, S., & Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and Language*, *98*, 74–88.

Lee, Y., Lee, H., & Gordon, P. (2007). Linguistic complexity and information structure in Korean: Evidence from eye-tracking during reading☆. *Cognition*, *104*(3), 495–534.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343.

Lemhöfer, K., Schriefers, H., & Hanique, I. (2010). Native language effects in learning second-language grammatical gender: A training study. *Acta Psychologica*, *135*(2), 150–158.

Lemhöfer, K., Schriefers, H., & Indefrey, P. (2014). Idiosyncratic Grammars: Syntactic Processing in Second Language Comprehension Uses Subjective Feature Representations. *Journal of Cognitive Neuroscience*, *26*(7), 1428–1444.

Lemhöfer, K., Spalek, K., & Schriefers, H. (2008). Cross-language effects of grammatical gender in bilingual word recognition and production. *Journal of Memory and Language*, *59*(3), 312–330.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–1177.

Levy, R. (2011). Integrating surprisal and uncertain-input models in online sentence comprehension: formal techniques and empirical results. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, 1055–1065.

Lewis, A. G., & Bastiaansen, M. (2015). A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex*, *68*, 155–168.

Lewis, A. G., Wang, L., & Bastiaansen, M. (2015). Fast oscillatory dynamics during language comprehension: Unification versus maintenance and prediction? *Brain and Language*, *148*, 51–63.

Lewis, R. L., & Vasishth, S. (2005). An Activation-Based Model of Sentence Processing as Skilled Memory Retrieval. *Cognitive Science*, *29*(3), 375–419.

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, *10*(10), 447–454.

Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language*, *63*(4), 447–464.

Lin, C. C. (2008). The processing foundation of head-final relative clauses. *Language and*

Linguistics, *9*(4), 813–838.

Lisman, J. E., & Jensen, O. (2013). The theta-gamma neural code. *Neuron*, *77*(6), 1002-1016.

Luo, H., & Poeppel, D. (2007). Phase Patterns of Neuronal Responses Reliably Discriminate Speech in Human Auditory Cortex. *Neuron*, *54*(6), 1001–1010.

Luo, Y., Zhang, Y., Feng, X., & Zhou, X. (2010). Electroencephalogram oscillations differentiate semantic and prosodic processes during sentence reading. *Neuroscience*, *169*(2), 654–664.

Luu, P., & Tucker, D. M. (2001). Regulating action: alternating activation of midline frontal and motor cortical networks. *Clinical Neurophysiology*, *112*(7), 1295–1306.

Luu, P., Tucker, D. M., & Makeig, S. (2004). Frontal midline theta and the error-related negativity: neurophysiological mechanisms of action regulation. *Clinical Neurophysiology*, *115*(8), 1821–1835.

MacWhinney, B., & Pléh, C. (1988). The processing of restrictive relative clauses in Hungarian. *Cognition*, *29*(2), 95–141.

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

Maier, A., Adams, G. K., Aura, C., & Leopold, D. A. (2010). Distinct superficial and deep laminar domains of activity in the visual cortex during rest and stimulation. *Frontiers in Systems Neuroscience*, *4*, 1–11.

Mak, W. M., Vonk, W., & Schriefers, H. (2002). The Influence of Animacy on Relative Clause Processing. *Journal of Memory and Language*, *47*(1), 50–68.

Mak, W. M., Vonk, W., & Schriefers, H. (2006). Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language*, *54*(4), 466–490.

Mak, W. M., Vonk, W., & Schriefers, H. (2008). Discourse structure and relative clause processing. *Memory & Cognition*, *36*(1), 170–181.

Mak, W., Vonk, W., & Schriefers, H. (2006). Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language*, *54*(4), 466–490.

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in Cognitive Sciences*, *8*(5), 204–210.

Makeig, S., Jung, T. P., Bell, A. J., Ghahremani, D., & Sejnowski, T. J. (1997). Blind separation of auditory event-related brain responses into independent components. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(20), 10979–10984.

Makuuchi, M., & Friederici, A. D. (2013). Hierarchical functional connectivity between the core language system and the working memory system. *Cortex*, *49*(9), 2416–2423.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–90.

Marr, D. (1971). Simple Memory: A Theory for Archicortex. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *262*(841), 23-81.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419-457.

McDonald, J. L. (2000). Grammaticality judgments in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics*, *21*(3), 395–423.

McDuff, S. G. R., Frankel, H. C., & Norman, K. A. (2009). Multivoxel Pattern Analysis Reveals Increased Memory Targeting and Reduced Use of Retrieved Details during Single-Agenda Source Monitoring. *Journal of Neuroscience*, *29*(2), 508-516.

Mecklinger, A., Schriefers, H., Steinhauer, K., & Friederici, A. D. (1995). Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory & Cognition*, *23*(4), 477–494.

Meyer, L., Grigutsch, M., Schmuck, N., Gaston, P., & Friederici, A. D. (2015). Frontal–posterior theta oscillations reflect memory retrieval during sentence comprehension. *Cortex*, *71*, 205–218.

Meyer, L., Obleser, J., & Friederici, A. D. (2013). Left parietal alpha enhancement during working memory-intensive sentence processing. *Cortex*, *49*(3), 711–21.

Meyer, L., Obleser, J., Kiebel, S. J., & Friederici, A. D. (2012). Spatiotemporal dynamics of argument retrieval and reordering: An fMRI and EEG study on sentence processing. Frontiers in Psychology, 3, 1–13.

Mitra, P. P., & Pesaran, B. (1999). Analysis of dynamic brain imaging data. *Biophysical Journal*, *76*(2), 691–708.

Moreno, I., de Vega, M., & León, I. (2013). Understanding action language modulates oscillatory mu and beta rhythms in the same way as observing actions. *Brain and Cognition*, *82*(3), 236–242.

Moreno, I., de Vega, M., León, I., Bastiaansen, M. C. M., Lewis, A. G., & Magyari, L. (2015). Brain dynamics in the comprehension of action-related language. A time-frequency analysis of mu rhythms. *NeuroImage*, *109*, 50–62.

Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012). Explicit and Implicit Second Language Training Differentially Affect the Achievement of Native-like Brain Activation Patterns. *Journal of Cognitive Neuroscience*, *24*(4), 933–947.

Morillon, B., Liégeois-Chauvel, C., Arnal, L. H., Bénar, C., & Giraud, A. (2012). Asymmetric function of theta and gamma activity in syllable processing: an intra-cortical study. *Frontiers in Psychology*, *3*, 248.

Moscovitch, M. (2008). The hippocampus as a "stupid," domain-specific module: Implications for theories of recent and remote memory, and of imagination. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *62*(1), 62-79.

Münte, T. F., Heinze, H., & Mangun, G. R. (1993). Dissociation of Brain Activity Related to Syntactic and Semantic Aspects of Language. *Journal of Cognitive Neuroscience*, *5*(3), 335–344.

Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically Based Sentence Processing Classes: Evidence from Event-Related Brain Potentials. *Journal of Cognitive Neuroscience*, *3*(2), 151–165.

Nevins, A., Dillon, B., Malhotra, S., & Phillips, C. (2007). The role of feature-number and feature-type in processing Hindi verb agreement violations. *Brain Research*, *1164*, 81–94.

Nieuwland, M. S., & Van Berkum, J. J. A. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, *18*(7), 1098–1111.

Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review*, *110*(4), 611-646.

Nyberg, L., Habib, R., McIntosh, A. R., & Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. *Proceedings of the National Academy of Sciences*, *97*(20), 11120–11124.

Onton, J., Delorme, A., & Makeig, S. (2005). Frontal midline EEG dynamics during working memory. *NeuroImage*, *27*(2), 341–356.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, *2011*, 1–9.

Orgassa, A., & Weerman, F. (2008). Dutch gender in specific language impairment and second language acquisition. *Second Language Research*, *24*(3), 333–364.

Osterhout, L. (1995). Event-Related Brain Potentials Elicited by Failure to Agree. *Journal of Memory and Language*.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), 785–806.

Osterhout, L., & Holcomb, P. J. (1993). Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes*, *8*(4), 413–437.

Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 786–803.

Otten, M., & van Berkum, J. J. A. (2008). Discourse-Based Word Anticipation During Language Processing: Prediction or Priming? *Discourse Processes*, *45*(6), 464-496.

Paczynski, M., & Kuperberg, G. R. (2011). Electrophysiological evidence for use of the animacy hierarchy, but not thematic role assignment, during verb-argument processing. Language and Cognitive Processes, 26(9), 1402–1456.

Peña, M., & Melloni, L. (2012). Brain oscillations during spoken sentence processing. *Journal of Cognitive Neuroscience*, *24*(5), 1149–1164.

Penolazzi, B., Angrilli, A., & Job, R. (2009). Gamma EEG activity induced by semantic violation during sentence reading. Neuroscience Letters, 465(1), 74–78.

Pérez, A., Molinaro, N., Mancini, S., Barraza, P., & Carreiras, M. (2012). Oscillatory dynamics related to the Unagreement pattern in Spanish. *Neuropsychologia*, *50*(11), 2584–2597.

Petersen, S. E., & Posner, M. I. (2012). The Attention System of the Human Brain: 20 Years After. *Annual Review of Neuroscience*, *35*(1), 73–89.

Pfurtscheller, G., & Lopes da Silva, F. (1999). *Event-related desynchronization. Handbook of Electroencephalography and Clinical Neurophysiology, revised series (Vol. 6)*. Amsterdam:

Elsevier B.V.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, *11*(3), 105-110.

Pickering, M. J., & Garrod, S. (2014). Self-, other-, and joint monitoring using forward models. *Frontiers in human neuroscience*, *8*, 132.

Polyn, S. M. (2005). Category-Specific Cortical Activity Precedes Retrieval During Memory Search. *Science*, *310*(5756), 1963-1966.

Pulvermüller, F., Eulitz, C., Pantev, C., Mohr, B., Feige, B., Lutzenberger, W., … Birbaumer, N. (1996). High-frequency cortical responses reflect lexical processing: an MEG study. *Electroencephalography and Clinical Neurophysiology*, *98*(1), 76–85.

Reali, F., & Christiansen, M. H. (2007). Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language*, *57*(1), 1–23.

Regel, S., Meyer, L., & Gunter, T. C. (2014). Distinguishing Neurocognitive Processes Reflected by P600 Effects: Evidence from ERPs and Neural Oscillations. *PLoS ONE*, *9*(5), e96840.

Roelofs, A. (2014). A dorsal-pathway account of aphasic language production: The WEAVER++/ARC model. *Cortex*, *59*(1874), 33–48.

Rogers, M. (1987). Learners Difficulties with Grammatical Gender in German as a Foreign Language. *Applied Linguistics*, *8*(1), 48–74.

Rommers, J., Dijkstra, T., & Bastiaansen, M. (2013). Context-dependent semantic processing in the human brain: Evidence from idiom comprehension. Journal of Cognitive Neuroscience, 762–776.

Roopun, A. K., Kramer, M. A., Carracedo, L. M., Kaiser, M., Davies, C. H., Traub, R. D., ... & Whittington, M. A. (2008). Period concatenation underlies interactions between gamma and beta rhythms in neocortex. *Frontiers in cellular neuroscience*, *2*, 1.

Roopun, A. K., Middleton, S. J., Cunningham, M. O., LeBeau, F. E. N., Bibbig, A., Whittington, M. A, & Traub, R. D. (2006). A beta2-frequency (20-30 Hz) oscillation in nonsynaptic networks of somatosensory cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(42), 15646–15650.

Rossi, S., Gugler, M. F., Hahne, A., & Friederici, A. D. (2005). When word category information encounters morphosyntax: An ERP study. *Neuroscience Letters*, *384*(3), 228–233.

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in*

Cognitive Sciences, *11*(6), 251-257.

Sassenhagen, J., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2014). The P600-as-P3 hypothesis revisited: Single-trial analyses reveal that the late EEG positivity following linguistically deviant material is reaction time aligned. *Brain and Language*, *137*, 29–39.

Scherag, A. (2004). The effects of late acquisition of L2 and the consequences of immigration on L1 for semantic and morpho-syntactic language aspects. *Cognition*, *93*(3), B97–B108.

Schriefers, H., Friederici, A. D., & Kuhn, K. (1995). The Processing of Locally Ambiguous Relative Clauses in German. *Journal of Memory and Language*, *34*(4), 499–520.

Service, E., Helenius, P., Maury, S., & Salmelin, R. (2007). Localization of Syntactic and Semantic Brain Responses using Magnetoencephalography. *Journal of Cognitive Neuroscience*, *19*(7), 1193–1205.

Singer, W. (1993). Synchronization of cortical activity and its putative role in information processing. *Annual Review of Physiology*, *55*, 349–374.

Singer, W. (2011). Dynamic formation of functional networks by synchronization. *Neuron*, *69*(2), 191–193.

St. George, M., Mannes, S., & Hoffinan, J. E. (1994). Global Semantic Expectancy and Language Comprehension. Journal of Cognitive Neuroscience, 6(1), 70–83.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137-149.

Steinhauer, K., White, E. J., & Drury, J. E. (2009). Temporal dynamics of late second language acquisition: evidence from event-related brain potentials. *Second Language Research*, *25*(1), 13–41.

Stolk, A., Todorovic, A., Schoffelen, J., & Oostenveld, R. (2013). Online and offline tools for head movement compensation in MEG. *NeuroImage*, *68*, 39–48.

Strauß, A., Kotz, S. A., Scharinger, M., & Obleser, J. (2014). Alpha and theta brain oscillations index dissociable processes in spoken word recognition. *NeuroImage*, *97*, 387–395.

Szewczyk, J. M., & Schriefers, H. (2013). Prediction in language comprehension beyond specific words: An ERP study on sentence comprehension in Polish. *Journal of Memory and Language*, *68*(4), 297-314.

Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Sciences*, *3*(4), 151–162.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *The Journal of Neuroscience*, *16*(13), 4240-4249.

Tallon-Baudry, C., Kreiter, A., & Bertrand, O. (1999). Sustained and transient oscillatory responses in the gamma and beta bands in a visual short-term memory task in humans. *Visual Neuroscience*, *16*(3), 449–459.

Tesche, C. D., & Karhu, J. (2000). Theta oscillations index human hippocampal activation during a working memory task. *Proceedings of the National Academy of Sciences*, *97*(2), 919–924.

Teyler, T. J., & DiScenna, P. (1986). The hippocampal memory indexing theory. *Behavioral Neuroscience*, *100*(2), 147-154.

Tokowicz, N., Kroll, J. F., de Groot, A. M. B., & van Hell, J. G. (2002). Number-of-translation norms for Dutch—English translation pairs: A new tool for examining language production. *Behavior Research Methods, Instruments, & Computers*, *34*(3), 435-451.

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing Subject and Object Relative Clauses: Evidence from Eye Movements. *Journal of Memory and Language*, *47*(1), 69–90.

Traxler, M. J., Williams, R. S., Blozis, S. A., & Morris, R. K. (2005). Working memory, animacy, and verb class in the processing of relative clauses. *Journal of Memory and Language*, *53*(2), 204–224.

Tulving, E., Kapur, S., Markowitsch, H. J., Craik, F. I., Habib, R., & Houle, S. (1994). Neuroanatomical correlates of retrieval in episodic memory: auditory sentence recognition. *Proceedings of the National Academy of Sciences*, *91*(6), 2012-2015.

Tulving, E., Voi, M. E. L., Routh, D. A., & Loftus, E. (1983). Ecphoric Processes in Episodic Memory [and Discussion]. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *302*(1110), 361-371.

Tyler, L. K., Marslen-Wilson, W. D., Randall, B., Wright, P., Devereux, B. J., Zhuang, J., Papoutsi, M., & Stamatakis, E. A. (2011). Left inferior frontal cortex and syntax: function, structure and behaviour in patients with left hemisphere damage. Brain, 134(2), 415–431.

Ueno, M., & Garnsey, S. M. (2008). An ERP study of the processing of subject and object relative clauses in Japanese. *Language and Cognitive Processes*, *23*(5), 646–688.

Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition*, *92*(1-2), 231-270.

Vaidya, C. J., Zhao, M., Desmond, J. E., & Gabrieli, J. D. . (2002). Evidence for cortical encoding specificity in episodic memory: memory-induced re-activation of picture processing areas. *Neuropsychologia*, *40*(12), 2136-2143.

van Ackeren, M. J., Schneider, T. R., Musch, K., & Rueschemeyer, S.-A. (2014). Oscillatory Neuronal Activity Reflects Lexical-Semantic Feature Integration within and across Sensory Modalities in Distributed Cortical Networks. *Journal of Neuroscience*, *34*(43),

van Berkum, J. J. A. (2012). The electrophysiology of discourse and conversation. In M. Spivey, K. McRae, & M. Joanisse (Eds.), The Cambridge Handbook of Psycholinguistics (pp. 589-612). Cambridge, UK: Cambridge University Press.

van Berkum, J. J. A, Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: evidence from ERPs and reading times. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*(3), 443-467.

van Berkum, J. J. A., Koornneef, A. W., Otten, M., & Nieuwland, M. S. (2007). Establishing reference in language comprehension: An electrophysiological perspective. Brain Research, 1146(1), 158–171.

van Berkum, J. J. A., Zwitserlood, P., Bastiaansen, M. C. M., Brown, C. & Hagoort, P., 2004. So who's ‘‘he’’ anyway? Differential ERP and ERSP effects of referential success, ambiguity and failure during spoken language comprehension. Paper presented at the Annual Meeting of the Cognitive Neuroscience Society, San Francisco.

van de Meerendonk, N., Kolk, H. H. J., Vissers, C. T. W. M., & Chwilla, D. J. (2010). Monitoring in Language Perception: Mild and Strong Conflicts Elicit Different ERP Patterns. Journal of Cognitive Neuroscience, 22(1), 67–82.

Van Dyke, J. A., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, *49*(3), 285–316.

van Elk, M., van Schie, H. T., Zwaan, R. A., & Bekkering, H. (2010). The functional role of motor activation in language processing: motor cortical oscillations support lexical-semantic retrieval. *NeuroImage*, *50*(2), 665–677.

Varela, F., Lachaux, J., Rodriguez, E., & Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, *2*, 229–239.

Vogel, E. K., & Luck, S. J. (2000). The visual N1 component as an index of a discrimination process. Psychophysiology, 37(2), 190–203.

Vos, S. H., Gunter, T. C., Kolk, H. H., & Mulder, G. (2001). Working memory constraints on syntactic processing: an electrophysiological investigation. *Psychophysiology*, *38*(1), 41–63.

Wang, X. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological Reviews*, *90*, 1195–1268.

Wang, L., Jensen, O., van den Brink, D., Weder, N., Schoffelen, J., Magyari, L., Hagoort, P., & Bastiaansen, M. C. M. (2012). Beta oscillations relate to the N400m during language comprehension. *Human Brain Mapping*, *33*(12), 2898–2912.

Wang, L., Zhu, Z., & Bastiaansen, M. (2012). Integration or predictability? A further specification of the functional role of gamma oscillations in language comprehension. Frontiers in Psychology, 3(187).

Wanner, E., & Maratsos, M. P. (1978). An ATN approach to comprehension. In M. Halle, J. Bresnan, & G. Miller (Eds.), *Linguistic Theory and Psychological Reality* (pp. 119–161). Cambridge, MA: MIT Press.

Warren, T., & Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition*, *85*(1), 79–112.

Waters, G. S., & Caplan, D. (1996). The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992). *Psychological Review*, *103*(4), 761–772.

Weiss, S., & Mueller, H. M. (2003). The contribution of EEG coherence to the investigation of language. Brain and Language, 85(2), 325–343.

Weiss, S., & Mueller, H. M. (2012). "Too Many betas do not Spoil the Broth": The Role of Beta Brain Oscillations in Language Processing. *Frontiers in Psychology*, *3*, 201.

Weiss, S., Mueller, H. M., Schack, B., King, J. W., Kutas, M., & Rappelsberger, P. (2005). Increased neuronal communication accompanying sentence comprehension. *International Journal of Psychophysiology*, *57*(2), 129–141.

Wells, J., Christiansen, M., Race, D., Acheson, D., & MacDonald, M. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology*, *58*(2), 250–271.

Wheeler, M. E., & Buckner, R. L. (2004). Functional-anatomic correlates of remembering and knowing. *NeuroImage*, *21*(4), 1337-1349.

Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences*, *97*(20), 11125-11129.

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: an event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, *16*(7), 1272-1288.

Wimber, M., Maaß, A., Staudigl, T., Richardson-Klavehn, A., & Hanslmayr, S. (2012). Rapid memory reactivation revealed by oscillatory entrainment. *Current Biology*, *22*(16), 1482-1486.

Womelsdorf, T., Schoffelen, J., Oostenveld, R., Singer, W., Desimone, R., Engel, A. K., & Fries, P. (2007). Modulation of neuronal interactions through neuronal synchronization. *Science*, *316*(5831), 1609–1612.

Woodruff, C. C., Johnson, J. D., Uncapher, M. R., & Rugg, M. D. (2005). Content-specificity of the neural correlates of recollection. *Neuropsychologia*, *43*(7), 1022-1032.

Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: A review. *Psychological Bulletin*, *133*(5), 800-832.

Nederlandse Samenvatting

Bijna iedereen leest elke dag wel een stuk tekst, terwijl dit een behoorlijk ingewikkelde taak is. Om te begrijpen wat we lezen moeten we representaties op verschillende niveaus zien te combineren. Denk hierbij aan klanken in een woord (fonologie) of de woorden in een zin (syntaxis). Alleen de correcte combinatie van deze hiërarchisch gestructureerde representaties zorgt voor een goede interpretatie. De verwerking van deze representaties wordt mogelijk gemaakt door complexe interacties tussen verschillende hersengebieden. Deze interacties zijn dynamisch en flucturen gedurende het lezen van een woord, een zin of een tekst. In dit proefschrift heb ik onderzocht of het lezen van zinnen en korte teksten wordt ondersteund door neurale oscillaties met een bèta-frequentie (neuronen geven signalen door in een bepaald ritme, het bèta-ritme is tussen de 13 en 30 Hz).

Er zijn momenteel twee concurrerende theorieën wat betreft de rol van deze bèta-oscillaties in taalbegrip. Ten eerste is er de *beta-syntax* theorie die veronderstelt dat bèta-oscillaties nodig zijn om de syntactische representaties op een gestructureerde manier te combineren tijdens zinsbegrip. Deze theorie voorspelt dat bèta-oscillaties toenemen als de zinsbouw complexer wordt, en dat bèta afneemt als syntactische verwerking wordt verstoord (bijvoorbeeld door een ongrammaticaal woord). De *beta-maintenance* theorie gaat er daarentegen van uit dat bèta-oscillaties een algemene rol vervullen en dus niet alleen belangrijk zijn op het niveau van zinsbouw, maar ook een rol spelen bij het vormen van een zinsbetekenis aan de hand van andere niveaus van linguïstische en extra-linguïstische informatie. Bèta-oscillaties reflecteren in dit geval de instandhouding of de verandering van het netwerk van de hersengebieden die betrokken zijn bij de constructie en representatie van de zinsbetekenis. Deze theorie voorspelt dat bèta toeneemt zolang de huidige zinsbetekenis instand moet worden gehouden, maar dat bèta afneemt zodra er wordt geanticipeerd dat de zinsbetekenis moeten worden aangepast aan de hand van linguïstische aanwijzingen (bijv. ongrammaticaal taalgebruik).

De eerste twee empirisiche hoofdstukken van dit proefschrift richten zich op aspecten van de *beta-syntax* theorie die nog niet eerder zijn onderzocht. In *hoofdstuk 2* wordt getest of informatie op tekst- of verhaalniveau (in dit geval de samenhang in betekenis) de bèta-oscillaties kan beinvloeden, die gerelateerd aan de verwerking van syntaxis op zinsniveau. Dit lijkt inderdaad het geval te zijn aangezien individuele woorden in zinnen met een coherente betekenis gepaard gingen met meer bèta-activiteit dan woorden in incoherente zinnen. Het was echter niet mogelijk om te bepalen welk mechanisme hiervoor verantwoordelijk was. Het kan inderdaad zijn dat dit verschil in bèta-activiteit onstaat omdat informatie op tekstniveau de verwerking van de zinsbouw beïnvloedt, maar het is ook mogelijk dat een ander aspect van zinsverwerking wordt beïnvloed (bijvoorbeeld de integratie van betekenis). In *hoofdstuk 3* heb ik gekeken of bèta-oscillaties eenzelfde rol spelen in een taal die pas op latere leeftijd is geleerd vergeleken met de moedertaal. In deze studie werden moedertaalsprekers van het Nederlands vergeleken met sprekers met Duits als moedertaal en Nederlands als tweede taal geleerd op latere leeftijd. In het bijzonder heb ik gekeken naar het effect van de verwerking van grammaticaal geslacht (denk aan de/*het fiets) op bèta-oscillaties. In beide groepen was de bèta-activeit lager bij een fout in grammaticaal geslacht dan bij een goede combinatie. Voor de groep met Nederlands als tweede taal was dit alleen het geval als ze gevraagd werden om nadrukkelijk op de grammaticale informatie te letten, en alleen als hun eigen specifieke representatie van grammaticaal geslacht werd meegenomen.

In *hoofdstuk 4* wordt er getest of het mogelijk is om woordinformatie een 'frequentie-label' te geven voordat iemand een zin leest. Zodoende zou de activiteit veroorzaakt door deze woordinformatie door middel van dit label te volgen zijn tijdens het lezen van een zin. Een eerder experiment heeft laten zien dat, als woorden worden aangeboden terwijl er een visuele stimulus flikkert in een bepaalde frequentie, deze frequentie later ook wordt gemeten als het woord wordt aangeboden zonder de visuele stimulus. Als dit inderdaad een robuuste manier is om woorden te binden aan een specifiek frequentie-label, dan kan dit worden gebruikt om te kijken wat er gebeurt tijdens het lezen van dat woord in een tekst. Helaas was het moeilijk om zo'n label te creëren, het effect lijkt niet robuust genoeg om daadwerkelijk te kunnen gebruiken om de activatie van een woord te volgen tijdens het lezen van een zin of tekst.

In *hoofdstuk 5* werden de twee theorieën betreffende de rol van bèta-oscillaties, de *beta-syntax* en de *beta-maintenance* hypothese, direct met elkaar vergeleken. Deze twee theorieën voorspellen tegenovergestelde resultaten als het gaat om bèta-oscillaties in tijdelijk-ambigue bijvoegelijke bijzinnen (bijv. Achteraf praat de vader, die de zonen bij het concert bewonderd hebben, met de dirigent over het optreden). Tijdelijk zijn er twee mogelijke interpreties van de bijzin (namelijk deze slaat op het subject van de hoofdzin, of dat de bijzin betrekking heeft op het object). Deze ambiguïteit wordt opgegeven zodra het woord komt dat de doorslag geeft welke van deze twee interpretaties juist is. De *beta-syntax* theorie voorspelt dat bèta-activiteit toeneemt op het moment dat dit beslissende woord wordt gelezen. Dit omdat de zin niet ongrammaticaal is, maar tegelijkertijd is een object-relatieve bijzin moeilijker te begrijpen dan een subject-relatieve bijzin. De *beta-maintenance* theorie voorspelt echter dat bèta-activiteit afneemt zodra het beslissende woord wordt gelezen, omdat de standaard interpretatie van een subject-relatieve bijzin moet worden aangepast tot een object-relatieve bijzin betekenis. Ik vond inderdaad deze reductie in bèta-oscillaties zodra het beslissende woord was gelezen, wat duidelijk ondersteunend bewijs levert voor de *beta-maintenance* theorie.

Al met al blijkt uit de bevindingen in dit proefschrift dat bèta-oscillaties bijhouden of de huidige zinsbetekenis moet worden vastgehouden of moet worden veranderd naarmate er meer informatie binnenkomt. Beta lijkt dus een algemene rol te vervullen, in tegenstelling tot wat de *beta-syntax* theorie voorspelt. De resultaten in dit proefschrijft zijn goed te integreren met andere onderzoeken die aantonen dat bèta niet alleen belangrijk is voor de zinsbouw of syntaxis bij het lezen van een zin, maar ook bijdraagt aan andere processen van zinsverwerking. Dit betekent niet dat men bèta niet meer kan gebruiken om syntaxis te onderzoeken, aangezien ook meerdere keren is aangetoond dat bèta varieert als zinsbouw verandert. Mijn onderzoek laat echter zien dat een bèta-modulatie tijdens lezen niet direct betekent dat de verwerking van de syntaxis in de text hiervoor verantwoordelijk is. Dit is juist een voordeel, het volgen van bèta-oscillaties geeft ons de mogelijkheid om te onderzoeken welke processen allemaal bijdragen aan het vormen van een betekenis op zins- of tekstniveau.

English Summary

Reading is a complex cognitive task that almost everyone engages in just about every day. Comprehending what we read requires the construction of hierarchically structured representations of the linguistic input at multiple levels of representation (syntax, semantics, phonology, etc.). Such processing is supported by the intricate interplay of a multitude of brain regions, that unfolds dynamically over the course of reading a word, a sentence, or a text. This thesis tackles the question of what can be learned about these interactions by examining neural oscillations in the beta (13-30 Hz) frequency range during the reading of sentences and short texts.

There are currently two main hypotheses about the role of beta oscillations during language comprehension. The 'beta-syntax' hypothesis claims that beta oscillations support the construction of structured syntactic representations during sentence comprehension. This entails that beta increases when such syntactic processing becomes more demanding and decreases when that processing is disrupted (e.g., due to a grammatical violation). The 'beta-maintenance' hypothesis takes a more domain-general approach, suggesting that beta oscillations during sentence comprehension reflect the maintenance or change of the underlying network of brain regions responsible for the construction and representation of the current sentence-level meaning. It does not limit the role of beta to the support of syntactic processing, but allows that beta may also reflect the influence of other types of linguistic and extra-linguistic information on the construction of a sentence-level meaning. This hypothesis suggests that beta increases whenever the current sentence-level meaning needs to be maintained under increased processing demands, and that beta decreases when the system anticipates (based on cues in the linguistic input; e.g., a grammatical violation or a semantic anomaly) that the current sentence-level meaning will need to change.

The first two empirical chapters of this thesis address aspects of the 'beta-syntax' hypothesis that have not yet been explored. In *Chapter 2* the question of whether discourse-level information (in this case semantic coherence) can influence beta oscillations related to sentence-level syntactic processing is assessed. Individual words in semantically coherent sentences exhibited higher beta power than those in incoherent sentences, suggesting that discourse-level information can indeed affect beta oscillations related to sentence-level processing. It was however not possible to unambiguously determine whether this beta difference was related

directly to the influence of discourse-level information on syntactic processing, or rather to its influence on some other aspect of sentence processing (e.g., semantic integration or lexical retrieval). In *Chapter 3* the question of whether beta oscillations are also related to syntactic processing in one's second language when that language has not been acquired from birth (or early childhood) was addressed. The effect of Dutch grammatical gender violations on beta oscillations was compared between native speakers of Dutch, and German late second-language learners of Dutch. Beta power was lower for grammatical violations in both native speakers and late second-language learners. For late-second language learners this was only the case when the task required participants to explicitly focus on grammatical information while reading the sentences, and only when trials were sorted according to participants' subjective representations of grammatical gender information.

In *Chapter 4* a first attempt was made to validate an approach that would allow lexical information to be assigned a frequency-specific 'tag' prior to reading a sentence and have its activation tracked as that sentence is read. The first step was to replicate an experiment demonstrating that the neural correlates of context information in the form of frequency-specific visual flicker during an encoding period while learning lists of words could be reinstated during a subsequent recognition period when the visual flicker was not present. If robust, this would provide the type of mechanism just described, enabling the assignment of frequency-specific 'tags' to lexical items (or different types of lexical information) in an encoding period that could then be tracked (reactivation of associated neural activity) during reading to see when that information becomes activated. Unfortunately, the results were mixed and it was decided that the effect was not robust enough to be relied upon for tracking lexical activation during reading.

A direct comparison of the 'beta-syntax' and the 'beta-maintenance' hypotheses was made in *Chapter 5*. The two hypotheses make opposing predictions when it comes to how beta oscillations should respond to temporarily ambiguous Dutch object-relative clause sentences (the less preferred and less expected construction type compared to subject-relative clause sentences) at the point of disambiguation. The 'beta-syntax' hypothesis predicts that beta power should increase when the sentence is disambiguated as an object-relative construction because syntactic processing, while not disrupted, does become more demanding. The 'beta-maintenance' hypothesis on the other hand predicts that beta power should decrease upon disambiguation towards an object-relative construction because the disambiguating element acts as a cue to the system that the current sentence-level representation (the system is assumed to construct the preferred subject-relative representation by default) is incorrect and needs to be revised. A clear beta power decrease after disambiguation was demonstrated, lending clear support to the 'beta-maintenance' hypothesis.

The overall conclusion is that beta oscillations appear to track the maintenance/change of the current sentence-level meaning under construction. This offers a more domain-general explanation than the 'beta-syntax' hypothesis, and accounts for the now numerous beta findings related to types of processing during sentence comprehension that are not strictly syntactic in nature. This does not however rule out the usefulness of beta oscillations for investigating various kinds of syntactic processing, as beta has now been extensively linked to a multitude of different types of syntactic manipulations. It simply means that not all modulations of beta oscillations during language comprehension necessarily indicate that syntactic processing is being affected by the experimental manipulation. At the same time, it open up the possibility to track the construction of sentence-level meaning construction and how that might be affected by different experimental manipulations, through the investigation of ongoing beta oscillations during reading.

Acknowledgements

I'm going to forget to mention some people. Some by accident. Others on purpose. You know where I live.

My time in Nijmegen has been invaluably enriched by the guidance of many great teachers, by the diversity of cultures, ethnicities, and ideas, and by the *gezelligheid* of all the wonderful people I have had the good fortune of encountering along the way.

First, I would like to thank my supervisors for all their patience, support, and guidance these last several years. Each of you has contributed uniquely to different aspects of my development as a scientist and an academic, and I hope you're mostly happy with the result because you'll have to live with this monster you've created for some time to come. Marcel, you were the obvious choice when I first decided I wanted to work on language and oscillations for my master's research project. That project was rough going, but the concept was novel and with a lot of time, effort, and frustration we ultimately managed to make something of it. I think what I appreciated most about your supervision was your calm rational approach to dealing with problems or disagreements, and your willingness to engage with ideas no matter how crazy, while providing a gentle nudge in the right direction. You were also excellent at ensuring that I was on the right track for the next steps in my career, and I'll always appreciate you giving me my first opportunity at publishing in international journals, and recommending me for the postdoc position in which I'm currently working. After I approached you about my master's project you immediately suggested that we talk to Herbert, and I'm extremely happy you did! This collaboration worked out perfectly and I'm very pleased we could extend it to my PhD. Herbert, your experience and wisdom have kept our various projects moving in the right direction. I have learned so much from you about (amongst many things) experimental methods, the importance of a clear and thorough description of those methods and of scientific results, and I think most importantly about honest reporting and scientific integrity. I would like to think I've always had a pretty good feeling for the latter, but you provided a shining example of how this should be approached in any scientific endeavor. I have always thoroughly enjoyed our (sometimes quite lengthy) meetings to discuss stimulus or experimental design, or the interpretation of results. Your sense of humor and insightful stories meant that the time always seemed to fly by. You have also always been very supportive, even after my PhD contract came to an end you did everything you could to ensure that I had an income, and at some point, even

offered me a place to stay! I appreciate everything, and I could not have asked for a better promotor. Jan-Mathijs, you came on board after I started my PhD and I think had no trouble fitting in with our little group (although knowing you the way I do now I think it probably took some time to get used to the slow, laid-back way in which we approached our meetings and discussions). You were the most hands-on of the supervisors, and I appreciate that you always made sure things were on track and progressing smoothly with our various projects. I have been very fortunate to learn about signal processing from one of the best in the business, and I was always very impressed with your ability to offer just the right level of explanation for my particular level of understanding at various stages of my development. You also taught me (whether intentionally or not) that there's usually a solution to most data analysis problems, and that if that solution is not readily implemented or available then you can go out and implement it yourself. I think this is one of the most valuable insights I will take away from my PhD work.

I worked with several people (both in and outside of Nijmegen) on projects alongside my PhD research and have very much enjoyed cultivating these relationships and learning from different styles and approaches to research. A very big thank you to all those I collaborated with during my time in Nijmegen: Christian, Diana, Ivan (and co.), Jona (and co.), Kristin, Lilla, Lin, and Tzvetan.

I am extremely grateful to the members of my reading committee, Prof. Peter Hagoort, Prof. Ole Jensen, and Prof. Dietmar Roehm, for all your time and effort, keeping beta high as you no doubt read diligently through each chapter.

To my paranymphs, you are two of my oldest (in Nijmegen anyway) and dearest friends and I'm incredibly happy and honoured to have you standing next to me during the defense. I appreciate all your help over the last weeks (months even) and I think it should be obvious that I wouldn't have been able to arrange this defense without you. Suus, we've known one another through the best of times and through the worst of times (since we worked in different buildings you could even call it a tale of two institutes). I appreciate many things, but most of all that you were always there when I needed to bend someone's ear (over a beer of course), and that you put up with me no matter what. We have a very similar sense of humour and a very similar outlook on life, and I hope we'll continue to be good friends in spite of that! Bart, I've always been so impressed by your logical, considered approach to just about everything. There are many things I appreciate about you too, but probably most of all that you manage to stay a true individual.

You always do what you want to do or what you think is the right thing to do, in spite of what other people might say or think. In combination with a logical, considered approach to life I think this makes for the best kinds of people. And of course, I also appreciate you always 'encouraging' me to join you for a *terrasbiertje* or two when the sun was shining, or for no particular reason other than that leaving work early is always a good idea!

I was incredibly lucky to be adopted by the Neurobiology of Language department at the MPI/DCCN (another big thank you to Peter for allowing me to be a part of the group) and I think the diversity of research and people has (amongst other things) both broadened my perspective on psycholinguistic and cognitive neuroscience research, and improved my ability to communicate my research to people from different backgrounds. Thank you to everyone in the group, whether we worked closely together or barely interacted. I also want to thank everyone with whom I've shared an office over the years. Alex, Bart, Jo, Koen, Ricarda, Linda, Maria, Peter, and Wendy, thank you all for putting up with my shenanigans and shenanignots.

There has also been excellent technical support throughout my research career in Nijmegen from many people in various ways, all of which I am extremely grateful for. Arno, Yvonne, and Ruud, you gave me my start in the CNS master's program and the support and friendly helpful attitudes you all displayed made the transition to the Dutch approach to things far easier for me. Els, Dirkje, and Kevin from the IMPRS, your enthusiasm and efficiency made the administrative side of my PhD far easier than it might have been and I don't think you get enough appreciation for the work that you do. I also especially enjoyed the pizza evenings and other social events, which I think are so important for PhD students, and you were always very good at encouraging these activities. Finally, a very big thank you to the admin (especially Tildie and Ina) and technical staff (especially Jessica, Paul, Marek, Uriel, and Mike) at the DCCN and the MPI for all your help and patience over the years.

My time in Nijmegen was not all work, I have met many strange, fascinating, exceptional, funny, talented, wonderful people in my time here. I won't be able to mention everyone by name but please all remember: I think you're pretty weird! I didn't play enough sports, there wasn't really time, but tennis was one I managed to keep up regularly thanks to Jonathan's and Maarten's enthusiasm and willingness to play at least twice a week most weeks. Thanks to all my other tennis partners in crime: Suus, Nicco, Bronagh, Hartmut, Alastair, and a few others every now and then, I really miss playing with you all. Volleyball was the other sport

I enjoyed once a week for a while, thanks to Jonathan, Sarah, Claudia, Martin, Marius, Svetlana, Michalis, Christian, Carmen, and a few others who were less consistent. I played basketball for a while but had to stop, mostly because I was pretty hopeless. Still, thanks to Jonathan, Flora, Tom, Malte, Marius, Michalis, Svetlana, Klodiana, Nietzsche, Christian, Carmen, Mark, Paul, and a few others for putting up with my running around, mostly like a headless chicken. Among my favourite social activities were Friday afternoon drinks, where there are far too many people to mention. Star Trek marathons were an excellent feature towards the end of my PhD, at the risk of TNG I'll say thanks Rick, Maarten, Alina, and sometimes one or two others. The pub quiz waxed and waned in popularity but our team tried to be relatively consistent. If only we could have been consistently in first place, right guys! Special thanks to the core team members over the years: Jeff, Daniel, Eric, Matthias, Lorijn, Amie; and then to others who made it all a lot of fun too: Gwilym, Elliot, Dan, Suus, Flora, Lianne, and many others I'm sure I've forgotten, including Henry and Adil who often helped us out instead of tending bar. To the people who've had the misfortune of being my housemate over the years (Suus, Wouter, and Jeff), I hope we can still be friends. Jeff, I'm glad you decided to follow me to Nijmegen, and I appreciate you being there through some really tough times. I hope the good times made up for the bad ones. People will know that I always like to mention that I have a bit of a philosophy background and I tried to keep up with the extreme thinking (think of philosophy as an extreme sport for the mind) as much as I could. Thanks to everyone at the monthly Foundations meetings and the various philosophy reading groups I attended for making that possible: Giulio, Frank, Ceci, Jolien, Eelke, Pim, Nicco, Lorijn, Jana, Peter, and many others. Kat, I could not leave you out of these acknowledgements and I'll always be grateful for all the good times.

To the gang and the gang++, you've made my time in Nijmegen more enjoyable than I could have imagined and I hope we never lose touch, and maybe even end up in some of the same places someday. Thank you all for being wonderfully weird people: Suus, Bart, Johanne, Martine, Nicco, Emma, Lisa, Maarten, Annika, Rick, Tobi, Lianne, Daniel, Gwilym, Elliot, Dan, Meryl, Sara, Ceci, Inge. We had a lot of fun together and I'll not soon forget it.

Finally, I want to thank my family and especially my sister Amy and my father Steve for putting up with me and for always being so loving and supportive. Amy, we haven't always got along and there are so many things we disagree about, but I love you anyway and I'm very proud of my little sister and all the wonderful work she does! Dad, I don't think I tell you enough how much

I appreciate you and everything you've done over the years for me. The time you somehow always made to help with homework or school projects, and your constant encouragement are the reason I turned into the inquisitive person I am today. There's no doubt that I would never have been able to pursue this dream of completing a PhD and becoming a scientist if it weren't for your love, understanding, and support. I am so grateful for everything.

Curriculum Vitae

Ashley Glen Lewis was born in 1985 in Durban, South Africa. He obtained a Bachelor of Arts degree in Cognitive Science from the University of KwaZulu-Natal in 2008 (cum laude), and completed a Bachelor of Arts Honours degree in Cognitive Science at the same university in 2009 (summa cum laude). Ashley was then awarded a Huygens scholarship to start a Master of Science research degree in Cognitive Neuroscience at Radboud University, Nijmegen in the Netherlands, which he completed in 2012 (bene meritum). Next, he started work as a PhD student at the Max Planck Institute for Psycholinguistics and the Center for Neuroimaging of the Donders Institute for Brain, Cognition and Behaviour, both in Nijmegen. His PhD work was funded by an International Max Planck Research School for Language Sciences Fellowship granted by the Max Planck Society. Ashley now lives in New Haven, Connecticut in the USA where he works as a postdoctoral research associate at Haskins Laboratories in the lab of Julie Van Dyke. His research is currently focused on individual differences in neural oscillations and ERPs during language comprehension.

Publications

Lewis, A. G., Schoffelen, J., Hoffmann, C., Bastiaansen, M. C. M., & Schriefers, H. (2016). Discourse-level semantic coherence influences beta oscillatory dynamics and the N400 during sentence comprehension. *Language, Cognition and Neuroscience*, Advance online publication.

Lewis, A. G., Lemhöfer, K., Schoffelen, J., & Schriefers, H. (2016). Gender agreement violations modulate beta oscillatory dynamics during sentence comprehension: A comparison of second language learners and native speakers. *Neuropsychologia*, *89*, 254-272.

Lewis, A. G., Schoffelen, J., Schriefers, H., & Bastiaansen, M. C. M. (2016). A predictive coding perspective on beta oscillations during sentence-level language comprehension. *Frontiers in human neuroscience*, *10*.

Lewis, A. G., & Bastiaansen, M. C. M. (2015). A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex*, *68*, 155-168.

Lewis, A. G., Wang, L., & Bastiaansen, M. C. M. (2015). Fast oscillatory dynamics during language comprehension: Unification versus maintenance and prediction? *Brain and language*, *148*, 51-63.

Moreno, I., De Vega, M., León, I., Bastiaansen, M. C. M., Lewis, A. G., & Magyari, L. (2015). Brain dynamics in the comprehension of action-related language. A time-frequency analysis of mu rhythms. *Neuroimage*, *109*, 50-62.

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 Glen Lewis*