

**Oscillatory neuronal dynamics
during lexical-semantic retrieval and
integration**

Markus Johannes van Ackeren

Doctor of Philosophy

The University of York

Psychology

August 2014

Abstract

Current models of language processing advocate that word meaning is partially stored in distributed modality-specific cortical networks. However, while much has been done to investigate where information is represented in the brain, the neuronal dynamics underlying how these networks communicate internally, and with each other are still poorly understood. For example, it is not clear how spatially distributed semantic content is integrated into a coherent conceptual representation. The current thesis investigates how perceptual semantic features are selected and integrated, using oscillatory neuronal dynamics. Cortical oscillations reflect synchronized activity in large neuronal populations that are associated with specific classes of network interactions. The first part of the present thesis addresses how perceptual semantic features are selected in long-term memory. Using electroencephalographic (EEG) recordings, it is demonstrated that retrieving perceptually more complex information is associated with a reduction in oscillatory power, which is in line with the information via desynchronization hypothesis, a recent neurophysiological model for long-term memory retrieval. The second, and third part address how distributed semantic content is integrated and coordinated in the brain. Behavioral evidence suggests that integrating two features of a target word (e.g., *Whistle*) during a dual property verification task, incurs an additional processing cost if features are from different (visual: *tiny*, audio: *loud*), rather than the same modality (visual: *tiny*, *silver*). Furthermore, EEG recordings reveal that integrating cross-modal feature pairs is associated with a more sustained low-frequency theta power increase in the left anterior temporal lobe (ATL). The ATL is thought to converge semantic content from different modalities. In line with this notion, ATL is shown to communicate with a widely distributed cortical network at the theta frequency. The fourth part of the thesis uses magnetoencephalographic (MEG) recordings to show that, while low frequency theta oscillations in left ATL are more sensitive to integrating features from different modalities, integrating two features from the same modality induces an early increase in high frequency gamma power in left ATL and modality-specific regions. These results are in line with a recent framework suggesting that local, and long-range network dynamics are reflected in different oscillatory frequencies. The fifth part demonstrates that the connection weights between left ATL and modality-specific regions at the theta frequency are modulated consistently with the content of the word (e.g., visual features enhance connectivity between left ATL and left inferior occipital cortex). The thesis concludes by embedding these results in the context of current neurocognitive models of semantic processing.

Table of Contents

Abstract	ii
Table of Contents	iii
List of Tables	vii
List of Illustrations	viii
Acknowledgements	xii
Author’s declaration	xiii
Chapter 1 Divide and conquer: Words are represented in a distributed cortical network	1
Representation of conceptual knowledge in the mind and brain	2
Links between language understanding and perception/action systems	3
A distributed cortical network of modality-specific regions supports language understanding	6
Cortical oscillations in cognition: Organization via synchronization	13
Physiological basis and measurement	13
A dynamic view on cognition	15
Oscillatory dynamics in memory and language	17
Thesis outline and aims	24
Chapter 2 : When less is more: Perceptual complexity of lexical-semantic content is reflected in reduced oscillatory neuronal activity	26
Abstract	26
Introduction	27
Methods	29
Participants	29
Experimental design	29
Analysis of event-related potentials	31
Analysis of spectral power	32
Results	33
Behavioral analysis	33
Event-related potentials.....	34
Spectral analysis	35
Discussion	36
Oscillatory power diminishes as information increases	37
Complex perceptual information is activated early during language processing.....	38

Relation to semantic richness.....	38
Conclusion.....	39
Chapter 3 : Cross-modal integration of lexical-semantic features during word processing: Evidence from oscillatory dynamics during EEG.....	41
Abstract.....	41
Introduction.....	42
Experiment 1.....	44
Methods.....	44
Participants.....	44
Stimulus material.....	45
Procedure.....	47
Results and Discussion.....	47
Experiment 2.....	49
Methods.....	50
Participants.....	50
Stimulus material.....	50
Procedure.....	50
Data recording and pre-processing.....	50
Spectral analysis.....	51
Statistical analysis.....	51
Source reconstruction.....	52
Connectivity analysis.....	52
Results and Discussion.....	53
General discussion.....	57
Integrating multimodal semantic information comes at a cost.....	58
Theta oscillations in left ATL during multimodal semantic feature integration.....	58
Relation to multisensory integration and cross-modal matching.....	59
Conclusions.....	61
Chapter 4 : Oscillatory dynamics in the theta band reflect cross-modal feature integration not switching.....	62
Abstract.....	62
Introduction.....	63
Methods.....	64
Participants.....	64
Stimulus material.....	65
Procedure.....	66

Data recording and pre-processing.....	67
Spectral analysis.....	67
Statistical analysis.....	68
Results	68
Behavioral results.....	68
Analysis of sustained theta power.....	69
Discussion	70
Chapter 5 : Oscillatory neuronal activity reflects lexical-semantic feature	
integration within and across sensory modalities in distributed cortical networks.	73
Abstract.....	73
Introduction	74
Materials and Methods	75
Participants.....	75
Experimental design.....	75
Data acquisition.....	77
Preprocessing	78
Time frequency analysis.....	78
Statistical analysis.....	78
Source reconstruction.....	79
Results	80
Behavioral analysis	80
Low frequencies are sensitive to semantic integration at the global scale.....	80
High frequencies reflect modality-specific semantic integration.....	81
Modality-specificity to semantic features in auditory and visual cortices.....	82
Discussion	83
Fast and slow components of lexical-semantic processing are reflected in the	
spectral and temporal profile of the signal.....	83
Oscillatory neuronal activity in the ATL reflects the distribution of lexical-semantic	
information in the cortex.....	84
Chapter 6 : Slow cortical oscillations connect left anterior temporal lobe to	
modality-specific cortical networks during semantic feature retrieval and	
integration	87
Abstract.....	87
Introduction	88
Participants.....	90
Experimental Design	91

Data acquisition.....	92
Data analysis and statistical testing.....	92
Results	94
Behavioral analysis	94
Seed-based functional connectivity analysis in left ATL.....	94
Discussion	95
Distributed semantic networks are linked via left ATL	96
Slow oscillations support long-range interactions during semantic retrieval	97
Possible confounds resulting from field spread	98
Conclusion.....	99
Chapter 7 : Towards a neurophysiological framework for semantic feature	
selection and integration.....	100
Summary of the main findings.....	101
Chapter 2	101
Chapter 3	101
Chapter 4	102
Chapter 5	102
Chapter 6	103
Selecting perceptual semantic features in memory	104
Timing of lexical access.....	105
Network dynamics during perceptual feature access	105
Integrating multimodal semantic content	107
Multimodal semantic content converges in left ATL.....	107
Oscillatory dynamics reflect local- and long-range network dynamics during	
semantic retrieval	109
Towards a neurophysiological model of semantic feature retrieval and	
integration	111
Language as a model for testing theories on cortical oscillations	112
Limitations	113
The time scale of oscillatory responses.....	113
The preparation period and pre-target integration.....	114
Recommendations for future research.....	114
Conclusions	115
Appendix	117
Bibliography	118

List of Tables

<i>Table 2.1. Behavioral results depicted as the mean entropy of semantic features in each of the three priming conditions.</i>	<i>34</i>
<i>Table 3.1. Matching of the experimental items. Scores were averaged over all items in each condition. P-values were computed using independent-samples t-tests. The standard error of the mean is provided in brackets.</i>	<i>46</i>
<i>Table 4.1. Matching of features for psycholinguistic variables and diagnosticity.</i>	<i>66</i>
<i>Table 4.2. Matching of nouns in the congruent and incongruent condition.</i>	<i>66</i>

List of Illustrations

- Figure 1.1. Gamma oscillations in visual feature binding. In the study by Tallon-Baudry and colleagues (1997) recognition of the Dalmatian dog among the monochrome pattern after the training phase elicited a larger power increase in the gamma band (~ 40 Hz). (From Tallon-Baudry et al., 1997)* _____ 16
- Figure 2.1. Experimental design. The trial started with a period of rest (1000 ms), which served as baseline for subsequent spectral analysis. The baseline period was followed by a modality-cue (Visual, Auditory, Haptic) (500 ms), a delay period (500 ms), and the target noun (1500 ms). After each trial, participants were asked to report two features of the target word from the cued modality.* _____ 30
- Figure 2.2. Global root mean square amplitude. RMS was calculated across all conditions and electrodes. Local peaks were found at (I) ~110ms, (II) ~150ms, (III) ~230ms, and (IV) ~350ms. 20ms time windows of interest was chosen for the first three peaks (I, II, III), and a 100ms time window for the last peak (IV) (shaded bars).* _____ 32
- Figure 2.3. Event-related potentials during target presentation. The left panel shows the event-related time course for 8 representative electrodes during the presentation of the target in each of the three cue conditions (Visual: black; Auditory: blue; Haptic: red). The left panel depicts a magnified illustration of the event-related time course at electrode PO3. Shaded areas depict the early (220-240ms) and late (300-400ms) time window in which the clustering algorithm identified a significant interaction between conditions. The distribution of each effect is represented in the adjacent topographic scalp maps. Black dots depict channels that are part of the significant cluster (mean time courses were smoothed at 10Hz for illustration purposes).* _____ 35
- Figure 2.4. Spectral decomposition showing the effect of the cued modality. A. Time-frequency decomposition showing the effect of the cued modality (rectangle) on the target word averaged over all electrode locations in the cluster. The topographical representation at the right bottom depicts the spatial extend of the cluster (black dots). B. Time frequency representations for each of the three conditions individually. During the time period of interest (rectangle) there is a reduction in all conditions relative to baseline. C. Bar graphs show the average signal decrease in the beta band for all three conditions. The graph shows that the visual (black) and auditory (blue) conditions show a larger decrease in beta power as compared to the haptic condition (red). D The spectrogram shows the distribution of the effect of cued modality in the frequency domain during the time period of interest. The illustration shows that the effect reaches significant (depicted as dashed line) between 20 and 30 Hz with a peak at 26 Hz.* _____ 36

Figure 3.1. Experimental design of the dual property verification paradigm. A The top panel provides an overview of the design in which a target was either paired with a cross-modal (visual-haptic [VH; HV], visual-auditory [VA; AV], auditory-haptic [AH; HA]), or modality-specific feature pair (Visual [V], Auditory [A], Haptic [H]). The three modalities of interest were visual, haptic, and auditory. B The bottom panel depicts the time course of a single trial. All words are presented one after the other. Therefore, features can only be fully integrated when the target appears (e.g., WHISTLE). _____ 45

Figure 3.2. Mean of the modality ratings for visual, haptic, and auditory features. The three spider plots indicate the mean rating score (Lynott and Connell, 2009; van Dantzig et al., 2010) over all features in the each of the three modalities of interest (Visual, Haptic, and Auditory). _____ 47

Figure 3.3. Cross-modal integration costs in verification times. Bar graphs depict the mean verification time in the MS (Visual, Auditory, and Haptic), and CM condition (Visual-Auditory, Auditory-Haptic, Visual-Haptic). Error bars denote standard error of the mean (** $p < .001$; ** $p < .01$). _____ 49

Figure 3.4. Modulation in low frequency cortical oscillations for the target word in a cross-modal or modality-specific context. A The top panel shows time-frequency representations, averaged over all significant clusters. The first two panels show the grand average percent signal change with respect to the baseline. The third panel depicts the masked statistical difference between the two conditions in t -values. The contour plot reveals one significant cluster in the theta range (4-6 Hz). B The first two bottom panels depict the topography of the effect in each condition (4-6 Hz, peak at 750-850 ms) relative to baseline. The third panel signifies the statistical difference between conditions in t -values. Electrodes within significant clusters are marked with dots ($p = .002$, cluster-corrected) _____ 54

Figure 3.5. Time-frequency plots for each of the 6 ROI. The ROI were middle anterior (MA), left anterior (LA), right anterior (RA), middle posterior (MP), left posterior (LP), and right posterior (RP) electrodes. Time-frequency representations depict the statistical difference in t -values for the target word in the CM versus MS feature context. The contours indicate the peak of the cluster-corrected statistical difference ($p = .002$). 55

Figure 3.6. Source reconstruction and connectivity analysis. A Source reconstruction of the effect in the theta band, depicted as thresholded z -values, reveals peaks in left ATL and MOG. B Bar graphs show a significant increase in the number of connections between ATL and the rest of the brain in the early time window (0-500 ms). In the late time window (500-1000 ms), only the CM condition shows a significant increase in the number of connections relative to baseline. Error bars depict SEM C Results of the

whole-brain connectivity analysis, seeded in the ATL (white dot). Connectivity maps show the difference in absolute, z-transformed, imaginary coherence between each condition and the baseline. In the early time window both conditions show a strong increase in connectivity between the ATL and a widespread cortical network. In the second time window, only the cross-modal condition shows continuing network activity above baseline. _____ 56

Figure 4.1. Timeline of the dual feature verification task. _____ 67

Figure 4.2. Spectral analysis of the post-stimulus period. A. The left panel depicts the theta (4-6 Hz) time course during the presentation of the target word when participants activate congruent features from different (black) or the same (red) modality. The shaded area highlights the time window of interest. The start signals that there is a significant difference between conditions. The topography at the top illustrates the statistical contrast between the cross-modal and modality-specific condition. The right panel depicts the timecourse for the incongruent feature context. B. Bar graphs show the difference between the cross-modal and modality-specific feature context for congruent (black), and incongruent (grey) feature-target pairs. Error bars depict standard error of the mean. The stars highlight the significant interaction effect and the effect of modality-specificity for congruent feature-target combinations. The topography at the top represents the interaction effect. Black dots depict the electrode locations that were part of the significant cluster. C. Spectral decomposition in the time window of interest, and across all electrodes. A peak difference effect of modality-specificity in the theta range (6 Hz) for the congruent (black), but not the incongruent features-target pairs (grey). Dashed lines illustrate the significance threshold ($\alpha = .05$, two sided) _____ 70

Figure 5.1. Design of the experiment. A. The word bus is associated with visual and auditory features. Participants are required to verify features from either the same (visual: red-big) or different modalities (audio-visual: red-loud). B Spider plots show that auditory (green) and visual features (blue) are rated as predominantly auditory (A) and visual (V) rather than haptic (H), olfactory (O), or gustatory (G). C. Features were always presented before target words. _____ 77

Figure 5.2. Low-frequency oscillations in sensor- and source space. A. Total power changes in the CM and MS condition relative to a 500 ms fixation baseline. Statistical differences between conditions are shown in the rightmost panel averaged over the identified cluster. The box depicts the significant time-frequency range. B. Topographies showing relative signal change across magnetometers between 2-8 Hz and 580-1000 ms. Significant channels are marked as dots in the rightmost panel. C. Source reconstruction revealed peaks in the ATL and parietal lobe. _____ 81

Figure 5.3. High-frequency oscillations in sensor- and source space. A. Total power changes in the CM and MS condition relative to a 500 ms fixation baseline, and the statistical difference between conditions are shown. The box depicts the significant time-frequency range. B. Topographies showing relative signal change across magnetometers between 80-120 Hz and 150-350 ms. Significant channels are marked as dots in the rightmost panel. C. Source reconstruction revealed peaks in left ATL and medial frontal lobe. _____ 82

Figure 5.4. Whole brain contrasts for modality-specific auditory and visual feature contexts in the gamma range. The first two panels show the contrast between the auditory and visual condition versus baseline. The last panel depicts the contrast between the auditory and visual condition. Gamma power is enhanced in visual areas for both conditions, but only the auditory condition shows a peak in pSTS/BA22 (dashed circle). All contrasts are corrected at the cluster level ($p < .005$). _____ 83

Figure 6.1. Design of the dual property verification task. After a period of fixation, two feature words are presented followed by the target word. Participants are asked to respond on catch trials only. _____ 91

Figure 6.2. Seed-based connectivity analysis between left ATL, ventral visual (red), and temporal auditory (blue) regions. The activation map illustrates the t -contrast between modality-specific visual (e.g., green-round-APPLE) versus auditory (noisy-thumping-DRYER) feature integration ($* p < .05$; $† p < .08$). _____ 95

Figure 7.1. Illustration of the oscillatory dynamics during the presentation of a visual word. Low frequency-theta power shows a sustained increase, while alpha and beta power are suppressed. _____ 107

Figure 7.2. The extended hub-and-spokes model. The foundation of the model is the hub-and-spokes model in which modality-specific semantic networks (green, blue, orange, grey) are connected via a central hub in the ATL (red). The extended model includes parameters illustrating which oscillatory frequencies operate at different levels of the model. Long-range connections are reflected in low frequency theta oscillations, while local interactions (illustrated as recurrent connections) are dominated by high-frequency gamma oscillations. Beta suppression could reflect the readout of modality-specific feature content. _____ 112

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor Dr. Shirley-Ann Rüschemeyer, who has provided me with outstanding support throughout previous years. She has given me the freedom and means to delve into a brave new field, but never failed to guide me whenever I lost the path. I would also like to thank the other members of my thesis advisory panel, Professor Beth Jefferies, and Dr. Silvia Gennari for their valuable suggestions and encouragement throughout the project.

During a study visit at the University Medical Center in Hamburg-Eppendorf, I had the great pleasure of working with Dr. Kathrin Müsch and Dr. Till Schneider, to whom I am indebted for helping me develop the skills necessary to ask new questions about my data.

I would also like to thank Rebecca Millman, Michael Simpson, and Garreth Prendergast for valuable advice and help with the MEG data acquisition.

Finally, I would like to thank the following brilliant undergraduate students who helped me with the data collection: Yaryna Dmytriv, Ben Lawrance, Lucy Homer, and Rebecca Waters.

Author's declaration

The present thesis contains original work that was completed by the author, Markus van Ackeren, under the supervision of Dr. Shirley-Ann Rueschemeyer

The current research was supported by a university studentship from the University of York, and a study visit grant from the Experimental Psychology Society

Some of the empirical work presented in the current thesis has been published in the following peer-reviewed journals:

van Ackeren MJ, Rueschemeyer S-A (2014) Cross-modal integration of lexical-semantic features during word processing: Evidence from oscillatory dynamics during EEG. *PLoS ONE* 9(7): e101042. doi: 10.1371/journal.pone.0101042

van Ackeren MJ, Schneider TR, Müsch K, Rueschemeyer S (in press) Oscillatory neuronal activity reflects lexical-semantic feature integration within and across sensory modalities in distributed cortical networks. *The Journal of Neuroscience*.

Furthermore, data from multiple chapters has been presented at the following conferences:

Van Ackeren MJ, Schneider, TR, Müsch, K, Rueschemeyer, S-A (2014) Integrating lexical-semantic features at different cortical scales: A dissociation between theta and gamma oscillations in the anterior temporal lobe. Poster presented at 19th International conference on Biomagnetism, Halifax, CAN.

Van Ackeren MJ, Schneider, TR, Müsch, K, Rueschemeyer, S-A (2014) Integrating lexical-semantic features at different cortical scales: A dissociation between theta and gamma oscillations in the anterior temporal lobe. Poster presented at Workshop on concepts, actions, and objects, Rovereto, IT

Van Ackeren MJ, Schneider, TR, Müsch, K, Rueschemeyer, S-A (2014) Integrating lexical-semantic features at different cortical scales: A dissociation between theta and gamma oscillations in the anterior temporal lobe. Poster presented at 20th annual meeting of the Organization for Human Brain Mapping, Hamburg, GER.

Van Ackeren MJ, (2014) Integrating lexical-semantic features at different cortical scales. Oral presentation at MEG UK 2014, Nottingham, UK

Van Ackeren MJ (2013) Oscillatory dynamics in EEG during multi-modal conceptual integration. Poster presented at Workshop on concepts, actions, and objects, Rovereto, IT

Chapter 1

Divide and conquer: Words are represented in a distributed cortical network

“Knowledge is power”

Francis Bacon

What would the world look like if we could not compare it to our rich system of conceptual knowledge¹? We would not know how to use a cup, or drive a car. Even the mere distinction between an animal and an object would be alien to us. Successful adaptation to the environment is built on the foundation of this conceptual system. But beyond understanding what is right in front of us, we encode complex conceptual knowledge in the form of words, allowing us to preserve, share, and recombine information beyond the limits of space and time.

Embodied theories of language postulate that accessing word meaning engages a distributed cortical network including areas primarily associated with perception and action (Barsalou, 1999; Martin and Chao, 2001; Barsalou et al., 2003; Pulvermüller and Fadiga, 2010). Specifically, different parts of the network are thought to represent different conceptual features. For example, retrieving the concept *apple* will engage neural pathways that encode what the object looks like (e.g., green and round), tastes like (e.g., sweet), and sounds like (e.g., crunchy). Indeed, a plethora of behavioral and functional imaging studies over the past decade have provided compelling evidence that these pathways are, to some extent, shared with perception and action systems (González et al., 2006; Hauk, Johnsrude, Pulvermüller, & Pulvermüller, 2004; Kiefer, Sim, Herrnberger, Grothe, & Hoenig, 2008; Simmons et al., 2007). However, while much has been done to understand the spatial extent (Goldberg, Perfetti, & Schneider, 2006), and flexibility (Hoenig et al., 2008; van Dam et al., 2012) in these networks, as well as the timing of lexical access (Pulvermüller et al., 2009), little is known about the neurophysiological mechanisms of *how* word content is retrieved and integrated in a distributed cortical network.

¹ Throughout the thesis, the terms *conceptual* and *semantic* will be used interchangeably.

The present thesis explores how patterns of synchronized neuronal activity contribute to the retrieval and integration of word meaning. To this end, the thesis will focus on two main themes. The first theme addresses how and when perceptual semantic features are selected during lexical-semantic retrieval. The second theme explores how these lexical-semantic features are integrated at multiple cortical scales, that is, within and across different modalities. The next part of this chapter will discuss empirical evidence suggesting that word meaning is partially stored in distributed cortical networks. Furthermore, current neurocognitive models will be presented that explain how conceptual features are integrated into coherent conceptual representations. The subsequent section will focus on previous neurophysiological work relating oscillatory dynamics to language and memory processes. Lastly, the final section will outline the general framework of the thesis and provide a preview of the empirical work presented in the following chapters.

Representation of conceptual knowledge in the mind and brain

Understanding the nature of the human conceptual system is a problem that has fascinated scholars since ancient times (Aristotle et al., 1993). However, the question how our knowledge system is organized has regained interest with the cognitive revolution in the 1950s. Psychologists at the time considered the human mind to be analogous to a computer. Perception and action were considered inputs and outputs to a modular cognitive system that operates through abstract symbolic processes (Fodor, 1981). However, it soon became clear that this view could not explain how abstract representations in the conceptual system relate to perceptual experiences of objects in the real world. This was famously referred to as the *symbol grounding problem* (Harnad, 1990).

A possible solution to this problem has been advocated in the form of *embodied* theories of conceptual representation, and language in particular. These theories postulate that conceptual knowledge is stored in a perceptual format, which shares common neural pathways with perception and action systems (Barsalou, 1999; Martin and Chao, 2001; Barsalou et al., 2003; Pulvermüller and Fadiga, 2010). For example, retrieving the meaning of a word reactivates or simulates previous experiences with its referent (Barsalou, 1999). The grounding problem does not arise in embodied theories because information is thought to be stored in the same format in which it was originally experienced.

Over the last decade, behavioral and neuroimaging studies have provided ample empirical support for embodied theories of language. Specifically, behavioral studies

have shown that understanding word meaning interferes with perception and action systems in a modality-specific way (Creem & Proffitt, 2001; Dils & Boroditsky, 2010; Gentilucci & Gangitano, 1998; Glenberg & Kaschak, 2002; Richter & Zwaan, 2009; Zwaan, Stanfield, & Yaxley, 2002). Furthermore, neuroimaging studies have demonstrated that word meaning engages neural substrates that are partially shared with perception and action systems. (Hauk et al., 2004; Kiefer et al., 2008; Simmons et al., 2007). For example, reading a word like cup activates part of the system that navigates our hand to grasp the object (Rueschemeyer, van Rooij, Lindemann, Willems, & Bekkering, 2010). The next section will take a closer look at the behavioral and functional imaging evidence for embodied theories of language and point out which results are still inconclusive.

Links between language understanding and perception/action systems

Behavioral studies investigating the link between language processing and perception/action systems can be categorized into those who show a) interference effects (e.g., inhibition or facilitation of reaction times) or b) mechanistic similarities between accessing word meaning and perception/action (e.g., after effects, modulation in kinematics) (Gentilucci and Gangitano, 1998; Creem and Proffitt, 2001; Glenberg and Kaschak, 2002). Here, both types of research are discussed for word meaning relating to action-related, versus perceptual information.

Action-related word meaning. The main source of evidence for embodied theories of language comes from studies investigating the relationship between language understanding and action planning. For example, Creem and Proffitt (2001) asked participants to grasp objects in various orientations, while performing a semantic paired associates, or visuo-spatial task. The authors demonstrated that participants are less likely to pick up the object in a goal-directed way, that is, at the handle, while performing the semantic task. This suggests that using an object in a functionally meaningful way requires retrieval of semantic knowledge from a system that is at least partially shared with language processing. These results could not be attributed to task difficulty, as the authors demonstrated in a control experiment. Here the task to pick up an object was replaced with a standard visuo-motor task (visual pursuit tracking). In contrast to the goal-oriented task in the first experiment, Creem and Proffitt found selective interference effects with the visuo-spatial, but not the semantic task. Taken together, these results demonstrate that at some level, conceptual knowledge is shared between action planning, and language processing.

Other studies have corroborated these findings showing that conceptual knowledge is shared between language understanding and action planning even at the level of specific action features. For example, Gentilucci and Gangitano (1998) asked participants to grasp a rod with the labels ‘*long*’ or ‘*short*’ printed onto the visible surface. The authors demonstrated that basic kinematic parameters during the execution of the grasping action, such as peak acceleration, velocity and deceleration were modulated as if the rod was indeed physically different in size. The authors concluded that participants rapidly, and automatically retrieve word meaning, which interferes with movement execution at a conceptual level.

However, the relation between language understanding and action planning is not restricted to the word level. Glenberg and Kaschak (2002) asked participants to make sensibility judgments on sentences describing either a movement towards their body (*open the drawer*) or away from their body (*close the drawer*). The results indicated that participants are slower to process a sentence when the response is incongruent with the sentence meaning, demonstrating that language understanding and action share a common motor code, even at a sentential level.

Taken together, behavioral studies have shown that performing goal-directed actions and language understanding compete for resources in a shared conceptual system. Interference effects are observed at the level of specific action features (e.g., spatial direction, acceleration, grip aperture, etc.), and go beyond the level of single word.

Perceptual word meaning. Similar to action-related language, accessing perceptual word meaning has been shown to interfere selectively with perceptual processing. In this line of research, most studies have focused on shared visual features such as shape, color, and motion. (Dils & Boroditsky, 2010; Richter & Zwaan, 2009; Zwaan et al., 2002). For example, Richter and Zwaan (2009) presented participants with a color patch, followed by a word, and another color patch. The word could be a color word, a non-color word or a non-word. Furthermore, the second color patch could either match the first one exactly, or differ minimally. The task was to perform a lexical decision on the word and subsequently decide whether the second color patch was the same as the first. The authors found that lexical decision latency was selectively reduced for color words, if word meaning was congruent with the first color patch. In addition, responses were slower during color discrimination on the second color patch, if the word meaning on the same trial had been incongruent with the first patch. The study by Richter and Zwaan demonstrates that perceiving a color affects retrieval of color related word meaning and vice versa.

Shared conceptual representations between language understanding and visual processing have also been reported for visuo-spatial information. Specifically, Zwaan Stanfield and Yaxley (2002) showed that verbal information about the location of an object induces a prediction/simulation about the likely shape of the object, which interferes with subsequent visual processing. Participants read sentences describing an object and its location, followed by a visual image. The task was to decide whether the object had been mentioned in the previous sentence. The rationale for the experiment was that the shape of an object changes with its location. For example, an *eagle in the sky* evokes the image of a bird with outstretched wings, while an *eagle in the nest* evokes the image of a bird with closed wings. Analysis of response latencies revealed that participants responded faster if the image was congruent with the shape implied by the location of the object in the preceding sentence. These results suggest that language understanding induces retrieval of visuo-spatial features, which selectively guide visual processing in the real world.

A different way to study commonalities between language understanding and visual perception is to identify perceptual phenomena that also exist in language understanding. For example, a well-known phenomenon in the study of visual motion is the visual motion aftereffect. The motion after effect describes the phenomenon that observing a motion in one direction – such as watching a waterfall – induces the perception of an illusory motion in the opposite direction when observing a physically static or randomly moving scene. Dils and Boroditsky (2010) presented participants with stories describing motion in the upward or downward direction followed by a randomly moving dot pattern. The task was to indicate whether dots were moving upwards or downwards. The authors demonstrated that similar to visual perception, understanding stories denoting a motion in one direction, elicits a motion-after effect in the opposite direction. These results further show that retrieving motion-related conceptual features during language understanding, and perceiving visual motion operate through shared cognitive systems.

In sum, evidence from behavioral studies suggests that retrieving modality-specific conceptual features, at some level, competes with perception and action in a shared conceptual system. These studies range from simple interference paradigms (Glenberg and Kaschak, 2002; Zwaan et al., 2002) to fine grained analyses of the effect of language understanding on movement kinematics (Gentilucci and Gangitano, 1998), or visual adaptation (Dils and Boroditsky, 2010). However, behavioral studies alone do not satisfy the anatomical claim of embodied theories of language, which states that language understanding engages shared neural substrates with perception/action systems. In the next section, I will discuss evidence from neuroimaging studies

supporting the idea of shared neural substrates between language understanding and perception/action systems.

A distributed cortical network of modality-specific regions supports language understanding

The previous section has provided an overview of behavioral studies showing functional links between language understanding and perception/action systems. At the neural level, embodied theories of language predict that words are at least partially represented in a distributed network of areas including neural pathways for perception and action. For example, words associated with visual information should engage ventral visual areas, while words denoting actions should engage a fronto-parietal motor network. A large number of studies over the last years have tested these predictions using imaging methods such as functional magnetic resonance imaging (fMRI). The majority of these studies have compared neural activation patterns to words that are associated with one dominant modality (e.g., action words, color words, etc.), to activation patterns during action execution or perceptual processing.

The motor network. In line with behavioral studies, functional imaging research has mainly focused on the representation of action words. In a seminal study Hauk, Johnsrude and Pulvermüller (2004) compared activation patterns during passive reading of words denoting foot, hand, or mouth actions (e.g., *kick*, *pick*, and *lick*) to patterns of activation in the motor system when participants were asked to move each effector individually. The authors found that action-related word meaning recruits regions in primary and pre-motor cortices in a somatotopic fashion, similar to action execution. Complementary results have been reported at the sentence level, when participants listen to sentences denoting foot, hand or mouth actions (Tettamanti et al., 2005). However, several recent studies failed to show primary motor activation in response to action words (e.g., Postle, McMahon, Ashton, Meredith, & de Zubicaray, 2008). While the role of primary motor cortex is still debated, many studies have shown selective sensitivity to action words in regions involved in action planning and action understanding. These regions comprise pre-motor cortex (Postle et al., 2008) and the inferior parietal lobule (van Dam et al., 2010, 2012).

Conceptual knowledge about actions is not only encoded in action verbs, but also in nouns describing objects or tools, which require a particular type of manipulation (Martin and Chao, 2001; Rueschemeyer et al., 2010b). For example, Rueschemeyer and colleagues (2010) found that nouns describing tools, which require a particular type of action (e.g., *cup*) engage a fronto-parietal motor network more than objects that do not

require an action in order to be used (e.g., *bookend*). The same fronto-parietal network has also been demonstrated to be sensitive to functional object knowledge of visually presented objects (Beauchamp and Martin, 2007). In conclusion functional imaging studies have shown that action related word meaning (action verbs, tools) engages parts of the cortical motor network that are also involved in action planning and action execution.

Perceptual networks. In line with research on the representation of action words in the cortical motor networks, different perceptual pathways have been shown to respond selectively to perceptual features of word meaning. Specifically, shared neural pathways between perception and word retrieval have been demonstrated for visual, auditory, tactile, and olfactory systems (González et al., 2006; Kiefer et al., 2008; Rueschemeyer, Glenberg, Kaschak, Mueller, & Friederici, 2010; Simmons et al., 2007; van Dam et al., 2012). Some of this evidence will be discussed below.

In the previous section, evidence was presented, which showed that color perception affects color word processing and vice versa, suggesting that the two domains, at some level, compete for the same neural resources (Richter and Zwaan, 2009). This conclusion has been corroborated by functional imaging research showing that the two engage overlapping cortical pathways. (Simmons et al., 2007; van Dam et al., 2012). For example, Simmons and colleagues (2007) asked participants to perform a color perception task and a property verification task. In the property verification task, participants were asked to indicate whether a specific color or action was associated with an object (e.g. *TAXI - yellow*). The study showed that a region in the left posterior fusiform gyrus is sensitive both to the perception of chromatic contrast and the retrieval of color-specific semantic features from words.

Evidence for shared neural substrates with perceptual processing has also been reported for sentences describing motion. For example, Rueschemeyer et al. (2010) presented participants with sentences that described an object motion towards the observer, a motion away from the observer, or no motion at all. The authors found that area MT/ V5, which is predominantly associated with the perception of visual motion, is also sensitive to understanding sentences describing object movement. However, MT/V5 was only sensitive to object motion, which was directed towards the observer. Rueschemeyer et al. (2010) conclude that motion sensitive regions are also recruited by motion related language. However, the authors argue that the effect was likely subject to top-down influences, such as relevance to the observer.

Beyond the visual modality, other studies have reported evidence for shared neural substrates between language understanding and perception for auditory

information. For example, in a combined fMRI and EEG study, Kiefer and colleagues (2008) presented subjects with an auditory perception, and a lexical decision task. The critical words in the lexical decision task were strongly associated with auditory features (*telephone*). The authors found that a region in superior/middle temporal gyrus responds selectively to both meaningful sounds and words with strong acoustic associations. Furthermore, the analysis of event-related potentials (ERP) revealed that words with acoustic features activate auditory regions within the first 150ms after word onset.

There is also evidence for neural overlap between language and perception in less salient modalities. Specifically, Gonzalez and colleagues (2006) asked subjects to read words with strong (*cinnamon*) or weak olfactory associations (*coat*). In line with previous findings in other modalities, the authors found regions in the bilateral piriform gyrus, and the right amygdala that are more sensitive to words with strong olfactory associations. These regions have been associated with olfactory perception in previous studies. Again, this is evidence for shared neural processes underlying the retrieval of olfactory features during word processing and olfactory perception.

In sum, functional imaging studies in recent years have provided compelling evidence for shared anatomical substrates between language understanding and perception/action systems. These diverse findings suggest that word meaning engages distributed networks of modality-specific regions in the cortex. However, it should be noted that data from functional imaging is necessarily correlational and the fact that modality-specific networks are sensitive to word content does not mean that are also necessary for retrieving modality-specific word content (Mahon and Caramazza, 2008; Willems and Casasanto, 2011). Furthermore, patients with lesions resulting from stroke or neurodegenerative disease often report semantic deficit at a more general level (e.g., deficits for specific categories), that is, beyond a single modality (Damasio, 1989; Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996; Jefferies & Lambon Ralph, 2006). In the next section, I will discuss neurocognitive models, which have been used to address some of these issues.

Neurocognitive models of conceptual organization. Current neurocognitive models agree that accessing conceptual knowledge through verbal or non-verbal information engages distributed modality-specific cortical networks. However, the way in which modality-specific networks contribute to the retrieval of conceptual knowledge is still debated (Mahon & Caramazza, 2008; Rueschemeyer, Lindemann, van Elk, & Bekkering, 2009; Willems & Casasanto, 2011). For example, it is not clear whether conceptual knowledge is encoded in modality-specific networks only (*distributed model*), or whether it converges in one (*hub-and spokes model*), or multiple

(*convergence zones*) high-level association cortices that link multiple modality-specific networks.

Distributed models postulate that conceptual information is encoded in modality-specific regions and the connections between them (Martin and Chao, 2001; Pulvermüller and Fadiga, 2010). A key prediction of this framework is that damaging modality-specific networks will directly impair the retrieval of modality-specific word content. For example, lesions in motor cortices will not only impair action execution, but also the understanding of words denoting actions and tools. This hypothesis has been addressed in patients with focal lesions to modality-specific networks, as well as healthy individuals where neural activity was modulated locally through transcranial magnetic stimulation (TMS) (Neininger and Pulvermüller, 2003; Buccino et al., 2005; Pulvermüller et al., 2005a; Boulenger et al., 2008; Tomasino et al., 2008; Papeo et al., 2009; Bonner and Grossman, 2012)

For example, Boulenger and colleagues (2008) compared masked priming effects for verbs and nouns in a group of patients suffering from Parkinson's Disease (PD) that were either on, or off medication. PD is a motor disease that is associated with subcortical atrophy in the substantia nigra, as well as a deterioration of apical dendrites in the motor cortices. The authors demonstrated stronger priming effects for action verbs when participants were on versus off medication, suggesting that medication directly improving motor function also selectively improves the retrieval of action-related word meaning. In a different study Neininger and Pulvermüller (2003) presented action verbs and concrete nouns with dominant visual, or visual and motoric associations to a group of stroke patients. Patients showed damage to either right frontal, or right temporal-occipital regions. The task was a speeded lexical decision task. The results revealed that patients with frontal lesions around motor cortices performed less accurately for action verbs than for concrete nouns. In contrast, patients with temporal-occipital lesions were less accurate for nouns with strong visual associations, but not action verbs. These data suggest that damage in modality-specific regions selectively impairs retrieval of modality-specific word meaning. However, an inherent problem with the interpretation of these results is due to the fact that verbs and nouns usually differ in more than just semantic aspects. For example, the dissociation in Neininger and Pulvermüller (2003) could equally well be explained with differences in word category (Mahon and Caramazza, 2008). To prevent this objection, more recent studies have compared nouns associated with semantic features from one dominant modality. For example, Bonner and Grossman (2012) tested word recognition for words with strong auditory, visual, or action associations in a group of patients with a logopenic variant of primary progressive aphasia. Specifically, these patients show cortical lesions in auditory association areas.

In line with predictions from distributed models of conceptual representation, patients showed a selective deficit for nouns with auditory associations (e.g., *thunder*). Furthermore, the authors demonstrated that impaired recognition of sound-related words was correlated with grey matter atrophy in auditory cortical networks. In conclusion, neuropsychological studies in patients with focal brain lesions provide evidence that modality-specific networks could causally be involved in the representation of word meaning.

Further evidence for a causal role of modality-specific networks in word retrieval has also been demonstrated in healthy individuals through the use of TMS. (Buccino et al., 2005; Pulvermüller et al., 2005a; Tomasino et al., 2008; Papeo et al., 2009). For example, Pulvermüller and colleagues (Pulvermüller et al., 2005a) applied single pulse TMS to hand regions or arm regions on primary motor cortex while subjects performed a lexical decision task on verbs denoting hand or foot actions. The results showed that participants responded faster to words denoting hand actions when hand regions were stimulated. Conversely, stimulating foot regions improved processing of foot related words. Other studies have used TMS to measure motor evoked potentials (MEP's) while subjects read action verbs or sentences. However, it is very difficult to integrate these findings as some studies report facilitation of MEP's (Papeo et al., 2009) during action language processing, while others report suppression (Buccino et al., 2005).

In summary, studies in patients with focal brain regions have demonstrated that modality-specific cortical networks are causally linked to the representation of word meaning (Neininger and Pulvermüller, 2003; Boulenger et al., 2008; Bonner and Grossman, 2012), which is also partially supported by TMS research in healthy individuals (Buccino et al., 2005; Tomasino et al., 2008; Papeo et al., 2009). However, it is noteworthy that in all of these studies retrieval of word meaning was merely affected by focal lesions in modality-specific cortical networks, yet never fully disrupted. In addition stroke patients rarely show major deficits in retrieving word content of a particular modality. Rather, depending on the lesion foci, patients are severely impaired in retrieving multimodal content about a particular category (e.g., tools, animals, or persons). Thus, while conceptual knowledge may be partially represented in modality-specific cortical networks, it is unlikely to be exclusively represented in these networks as postulated by distributed models of conceptual representation (Martin and Chao, 2001).

In order to account for category-selective semantic deficits (e.g., tools, or animals) in stroke patients with focal lesions in temporal and parietal cortices, Damasio and others have argued that perceptual knowledge is integrated via multiple *convergence*

zones (Damasio, 1989; Damasio, Grabowski, & Tranel, 1996). The model argues that different convergence zones are sensitive to different sets of perceptual features. Damage to a convergence zone, following stroke, may result in selective deficits in categories that are most reliant on a particular set of features. For example, animals and food are more easily distinguished based on shape, color and motion, while knowledge about tools may depend more on visuo-motor associations. The idea of graded perceptual integration has successfully been demonstrated in connectionist models (Plaut, 2002), and remains an important aspect of more recent embodied theories of language (Meteyard & Vigliocco, 2008; Meteyard & Cuadrado, 2010; Vigliocco, Meteyard, Andrews, & Kousta, 2009). However, convergence zone models alone fail to explain the whole clinical spectrum of semantic deficits. For example, patients with semantic dementia (SD), a neurodegenerative disease affecting the anterior temporal lobes, show a generalized semantic deficit that cannot be reduced to a single category, or modality.

The hub and spokes model (Patterson et al., 2007), postulates that conceptual knowledge is partially stored in a distributed network of modality-specific areas, but linked through supramodal association cortices in the anterior portion of the temporal lobe. In other words, information from all modality-specific networks converges in a single semantic hub. In connectionist terms, the hub can be thought of as a hidden layer in which conceptual knowledge is clustered in multidimensional features space. This could explain, how individuals are able to associate objects that do not share a dominant perceptual feature. Indeed SD patients with damage the anterior temporal lobes are more likely to categorize objects in terms of perceptual similarity, which can result in undergeneralization for atypical members of a category (a penguin is a bird), or overgeneralization for perceptually similar non-members of a category (Lambon Ralph et al., 2010). In line with these observations in clinical populations, computational models have shown that a hub-like hidden layer is necessary to account for the ability to generalize and abstract beyond simple perceptual features (Rogers et al., 2004).

Taken together, current neurocognitive models on conceptual representation largely agree that perceptual word meaning is at least partially stored in distributed modality-specific cortical networks (Damasio, Grabowski, & Tranel, 1996; Patterson et al., 2007; Pulvermüller & Fadiga, 2010; Vigliocco et al., 2009). Neuropsychological evidence further suggests that modality-specific networks might be linked via one central hub (Patterson et al., 2007), or multiple convergence zones (Damasio, 1989; Damasio, Grabowski, & Tranel, 1996). However, there is currently a striking explanatory gap in all of these theories as to how modality-specific semantic content is

integrated into more complex multimodal representations. The next section will discuss this problem in more detail and outline a possible way to study conceptual integration.

Binding multimodal conceptual features. Previous research on conceptual representations has mainly focused on words associated with one dominant modality. Some of this evidence has been discussed in earlier sections of this chapter. For example, words associated with actions (e.g., *grasp*) engage a fronto-parietal motor network (Hauk et al., 2004; Postle et al., 2008), while words describing visual information (e.g., *red*) engage a ventral visual occipito-temporal network (Simmons et al., 2007). In contrast, other research has identified regions that are sensitive to a particular category, such as animals or tools (Damasio, Grabowski, & Tranel, 1996; Tranel, Damasio, & Damasio, 1997). However, the mechanism of how modality-specific features are integrated into complex multimodal object representations is still poorly understood.

Functional imaging studies have only recently started to address the question how multimodal conceptual information is retrieved and integrated. In a recent study, van Dam and colleagues (2012) presented participants with words that are strongly associated with bimodal (i.e., visual and action) conceptual features. For example, a *tennis ball* is strongly associated with visual information, (yellow, round) and information on how to use it (hit with racket). The authors demonstrated that words with bimodal associations engage both ventral visual networks (fusiform gyrus), as well as posterior motor networks (inferior parietal lobule). Additionally, the authors showed that modality specific changes in the BOLD signal were subject to top-down modulations. That is, activation in motor areas was enhanced when participants were asked what to do with the object versus what it looked like. The study illustrated what, up to that point, had only been an implicit assumption in neurocognitive models of language; word meaning engages multiple modality-specific networks simultaneously, which are modulated, as a function of what information is currently needed.

However, the idea that conceptual features from multiple modalities are activated simultaneously is not straightforward. Pecher Zeelenberg and Barsalou (2003) demonstrated that property verification latencies increase when the current property (BLENDER-loud) is from a different (CRANBERRIES-tart) versus the same (LEAVES-rustling) modality as the previous trial. The *conceptual* modality switch effect complements earlier findings showing *perceptual* switching costs in visual and auditory discrimination tasks (Spence, Nicholls, & Driver, 2001). Pecher and colleagues (2003) conclude that this is evidence for a distributed representation of word meaning in modality-specific conceptual networks. However, the study also shows that accessing

and switching between different modalities potentially recruits different, and more costly, mechanisms than accessing information from the same modality.

Taken together, few studies to date have attempted to bridge the gap between word content from a single dominant modality and words describing complex multimodal objects. A first step towards this direction was the study by van Dam et al. (2012) in which the authors illustrated that words strongly associated with features from two modalities engage modality-specific cortical networks relating to each feature type simultaneously. However, the fact that we access and bind information from multiple modalities simultaneously is hard to reconcile with the finding that evaluating conceptual features from different modalities in short succession incurs a processing cost (Pecher et al., 2003). Neurocognitive models presented in the previous section propose that conceptual knowledge from different modalities is integrated through a central hub (Patterson et al., 2007), or multiple convergence zones (Damasio et al., 2004). However, while these models can be used to generate predictions about *where* conceptual information might be integrated, they are silent with respect to *how* information is integrated within and across different modalities. In the study of visual perception, this is known as the binding problem (Treisman, 1996). Specifically, the problem describes how visual shape and color of an object are combined into a coherent visual percept of e.g. an apple. In recent years, evidence from neurophysiological studies in humans and animals has led to the hypothesis that information from different regions in the brain can be associated or integrated if the underlying neural populations discharge in synchrony (Engel, Fries, Singer, & others, 2001; Fries, 2009; Tallon-Baudry & Bertrand, 1999). In the next section, I will discuss how synchronous activity in large populations of neurons generates oscillatory activity at the level of the scalp, as well as evidence for the role of rhythmic neuronal activity in perceptual binding, memory, and language – three cognitive domains that are key to understanding the processes underlying conceptual knowledge representation.

Cortical oscillations in cognition: Organization via synchronization

Physiological basis and measurement

The ability to perceive the world in a coherent and stable manner arises from interactions between neural populations in the brain. At the microscopic level, information in the central nervous system is transmitted through neuronal axons and synaptic connections between neurons. Both of these mechanisms induce measurable electrical potentials. Information transmission through neuronal axons can be recorded through invasive techniques such as single cell or multi unit recording, which are

frequently applied in animal models. However, as these invasive techniques are not feasible in healthy human individuals, neurophysiological research in humans has been restricted to measuring electromagnetic brain activity at the surface of the scalp.

These macroscopic dynamics are the result of a chain of events at the microscopic level. When an action potential reaches the synapse, neurotransmitters are released into the synaptic cleft between two neurons. These neurotransmitters bind to receptors on the post-synaptic membrane. As a result, ion channels open or close, such that ions diffuse through the membrane on the receiving synapse. This modulation of ion influx induces a change in the post-synaptic membrane potential, which lasts an order of magnitude longer than an action potential. This is critical because it allows projections from multiple neurons to summate near the dendrite of the receiving neuron. In other words, synchronous firing of multiple neurons elicits a larger increase in the post-synaptic membrane potential. This causes a separation between a positive and a negative charge, which is called a dipole. Finally, the summation of spatially aligned dipoles from multiple post-synaptic potentials will generate a signal that is strong enough to be recorded at the level of the scalp. These rhythmic dynamics can be measured as electrical potentials (EEG) or magnetic fields (MEG) (Niedermeyer and Lopes da Silva, 1993).

In contrast to functional imaging (PET, fMRI), electrophysiological techniques (EEG and MEG) provide a very precise measure of the timing of neurophysiological responses. As a result, these techniques are widely used to study the latency, and spectral structure of a neural event. With respect to the latter, a line of research, beginning in the previous century (Berger, 1929), has associated oscillatory dynamics at different frequencies with distinct cognitive functions, ranging from low-level perceptual encoding to high-level memory and language (Donner & Siegel, 2011; Engel et al., 2001; Tallon-Baudry & Bertrand, 1999; von Stein & Sarnthein, 2000). For example, Hans Berger (1929), the pioneer of electroencephalography, observed a sustained cortical rhythm around 10 Hz at the back of the brain when the participant's eyes were closed. This rhythm disappeared when the participant's eyes were opened (Berger, 1929). Since its discovery, the alpha rhythm has been associated with a number of different cognitive functions such as attention, memory, and consciousness (Donner & Siegel, 2011; Fell & Axmacher, 2011; Jensen, Kaiser, & Lachaux, 2007; Klimesch, Freunberger, & Sauseng, 2010; Panagiotaropoulos, Deco, Kapoor, & Logothetis, 2012). It is now thought that alpha oscillations reflect a more general inhibitory, or gating mechanism that operates via thalamo-cortical pathways (Jensen & Mazaheri, 2010).

The current section has provided a primer of the physiological mechanisms underlying neuronal communication and how it can be measured at the macroscopic

scale using MEG and EEG. Through the example of the alpha rhythm, it has been argued that cortical oscillations are linked to specific cognitive processes. In recent years, other rhythms have been identified and associated with different cognitive functions. These are the delta rhythm (1-3 Hz), the theta rhythm (3-7 Hz), the beta rhythm (13-30 Hz) and the gamma rhythm (>30 Hz) (Niedermeyer and Lopes da Silva, 1993). These rhythms have been observed in different, or sometimes the same, cortical pathway (Donner and Siegel, 2011), and some have been linked to specific neurotransmitters (Singh, 2012). The next sections will review some of the evidence relating oscillatory dynamics at different frequencies to memory and language. But first, an influential theoretical framework will be presented that explains how oscillatory activity, or synchronous discharges in large neuronal populations, supports the integration of information at a larger scale.

A dynamic view on cognition

The human brain is unique in the way it allows us to adapt to the demands of an ever-changing environment. In line with this idea, the study of human cognition has undergone a major paradigm shift in recent years. In classical frameworks, perception is considered as the process of matching objects in the environment to context-invariant internal representations. In contrast, the *dynamicist view* of cognition emphasizes that cognitive processes are tuned to the agent's interaction with the environment (Engel et al., 2001). Accordingly, brain states should not be considered as stable representations, but rather as biases towards our interaction with the environment. However, this introduces the question how we are nevertheless able to experience coherent and stable objects in our environment.

A tentative solution to this problem lies in the fact that the perception of an object in the environment elicits synchronized, or correlated, discharges in distributed neural populations. According to the *temporal binding hypothesis* (Singer and Gray, 1995), these correlated discharges increase the impact, or saliency of a neural assembly within the network as a whole. This idea relates to the mechanics of synaptic transmission in which the number of synchronized neuronal discharges is proportional to the amplitude of the post-synaptic potential. In other words, the more neurons participate in synaptic transmission, the more likely the information will be transmitted further through the system. For example, neural populations encoding different features of a dog will frequently correlate with each other, creating a bias towards perceiving the dog as a coherent and stable percept.

An example of such a bias is a study by Tallon-Baudry and colleagues (Tallon-Baudry et al., 1997a), in which the authors recorded electrical potentials from the scalp

while presenting a modified version of the famous Dalmatian dog picture (Figure 1.1). Naive participants do not perceive the Dalmatian dog among scattered monochrome blobs. However, with some training the hidden dog is easily perceived by all participants. Thus, while the physical properties of the image are identical before and after training, the way in which the information was integrated or bound changed after the training phase. In line with predictions from the temporal binding hypothesis, the authors found an increase in oscillatory power around 40Hz when a participant consciously perceives the dog. The authors conclude that integrating visual object features is reflected in synchronized neuronal activity, or an increase in oscillatory power.

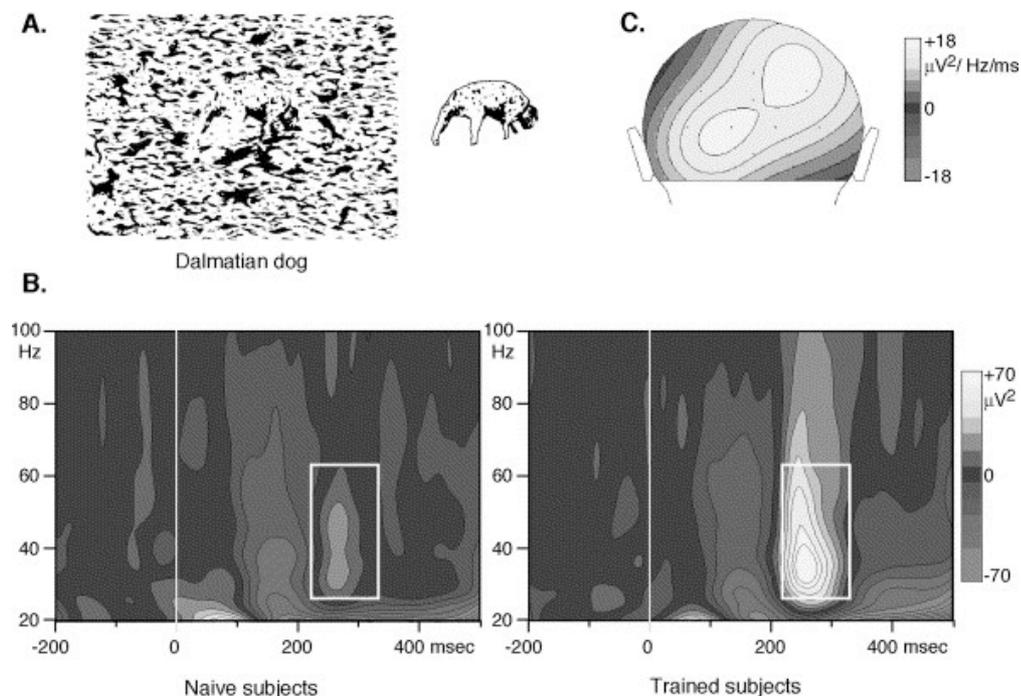


Figure 1.1. Gamma oscillations in visual feature binding. In the study by Tallon-Baudry and colleagues (1997) recognition of the Dalmatian dog among the monochrome pattern after the training phase elicited a larger power increase in the gamma band (~ 40 Hz). (From Tallon-Baudry et al., 1997)

The current section introduced the temporal binding hypothesis, which argues that synchronous neuronal activity could explain how information in the brain converges towards a coherent interpretation of the environment. This was demonstrated in humans, where successful integration of a coherent percept (Dalmatian dog) from visual noise was correlated with enhanced power in cortical gamma oscillations. The review of

embodied theories of language in earlier sections concludes with a related enigma: How is stored knowledge from distributed modality-specific networks integrated into a coherent conceptual representation? The next section first discusses the role of oscillatory dynamics for cognitive functions, and modalities other than vision. Subsequently, these data are used to generate a prediction for how conceptual features could be integrated into a coherent representation.

Oscillatory dynamics in memory and language

The study by Tallon-Baudry et al. (1997) has demonstrated that feature binding in the visual modality is reflected in high frequency oscillatory dynamics in the gamma range. In recent years, these findings have been corroborated in vision (Tallon-Baudry et al., 1997b; Tallon-Baudry and Bertrand, 1999; Donner and Siegel, 2011), and extended to other perceptual modalities (Kaiser et al., 2002; Cervenka et al., 2011), as well as the motor system (Donoghue et al., 1998; Muthukumaraswamy, 2010). A number of studies have also reported changes in lower frequency ranges, such as the alpha band. However, these slow oscillatory dynamics are often attributed to high-level cognitive functions, such as attention (Foxye, Simpson, & Ahlfors, 1998; Foxye & Snyder, 2011).

Oscillatory dynamics have also been observed during tasks involving higher cognitive functions such as decision making, attention, memory, and language (Donner & Siegel, 2011; Klimesch, 1999; Klimesch et al., 2001; Klimesch et al., 2010; Sauseng, Griesmayr, Freunberger, & Klimesch, 2010; Summerfield & Mangels, 2005; Wu, Chen, Li, Han, & Zhang, 2007). Of particular interest for the question how word meaning is integrated, are studies investigating the role of oscillatory dynamics in memory and language processing. The next two sections will provide an overview of research that has linked oscillations in distinct frequency ranges to memory maintenance and retrieval as well as lexical-semantic and sentence level language processing.

Long-term memory retrieval. The strongest evidence that oscillatory dynamics are involved in memory processes came from recordings of local field potentials in free moving rodents. In a seminal study, O'Keefe and Recce (1993) demonstrated that the spatial location of the rat is encoded within the phase of the theta rhythm. In humans, electrophysiological studies have reported changes in oscillatory activity in the theta (3-7), alpha (8-12 Hz), and gamma (>30 Hz) range during various memory-related tasks (Jensen & Tesche, 2002; Klimesch, Schimke, & Schwaiger, 1994; Lisman, 2010; Osipova et al., 2006; Raghavachari et al., 2001; Sauseng et al., 2010). For example, in an MEG study by Osipova and colleagues (2006), participants were asked to memorize a set of images depicting buildings or landscapes. The authors found that theta and gamma

power were enhanced during both retrieval and encoding of items that were correctly remembered versus items that were not remembered, or had not been presented in the training set. This effect is known in the literature as the subsequent memory effect (Brewer et al., 1998). Furthermore, the authors showed, using a beamformer source reconstruction method, that the effect in the gamma band was strongest around occipital cortex. Based on these results, the authors conclude that theta oscillations are involved in mnemonic operations, while gamma oscillations could reflect a top-down facilitation of representational content. The latter is in line with the idea presented in the previous section, suggesting that high-frequency gamma oscillations are involved in local binding of perceptual features (Tallon-Baudry and Bertrand, 1999).

Theta oscillations in particular have also been shown to be sensitive to the quality of a memory trace. In cognitive models of memory retrieval, the quality of a memory is captured in the distinction of familiarity (poor quality) versus recollection (high quality). A number of studies in previous years have demonstrated that theta power is enhanced during recollection versus familiarity (Guderian & Düzel, 2005; Klimesch et al., 2001; Klimesch, Schimke, & Schwaiger, 1994). For example, Klimesch et al. (2001) asked participants whether they consciously remembered seeing a word, which had been presented earlier, or whether the word was just familiar. Enhanced theta power was observed during items that were consciously remembered versus familiar. A more recent MEG study by Guderian and Düzel (2005) replicated this effect with images of faces, and showed, using source reconstruction, that the theta modulation was reflected in a widespread cortical network including temporal, frontal, and occipital regions. The authors conclude, that recollection engages a more distributed cortical network and theta oscillations in particular could be involved in binding this information into a coherent memory trace.

In addition to an increase in theta and gamma power, many studies report a stimulus-induced reduction or desynchronization in the alpha and beta band (8-30 Hz) during long-term memory retrieval (Burgess and Gruzelier, 2000; Spitzer et al., 2009; Khader and Rösler, 2011). For example, Khader and Rösler (2011) trained participants to associate abstract words with a) a spatial location on a checkerboard, b) an image of an object (cup), c) or both. During the test phase participants saw two words and were asked whether they shared a common features. Thus, the authors could manipulate parametrically the number of features participants were required to retrieve (2,3,4), as well as the information type (object identity, spatial location). Using EEG, Khader and Rösler demonstrated a parametric effect for theta synchronization as well as alpha desynchronization. However, only the latter showed a topographic interaction as a function of stimulus type. The authors conclude that theta oscillations might reflect more general

memory retrieval processes, while beta oscillations reflect the reactivation of stored information. In line with this finding, a recent framework, the information via desynchronization hypothesis (Hanslmayr et al., 2012) has proposed that desynchronization in the alpha/beta band (8-30 Hz) could be related to the richness, or entropy, of information that is currently represented in the system. In their model, the authors illustrate that desynchronization in neural population effectively increases their informational content. Thus, the richer the informational content of an item in our memory system, the more desynchronization the model predicts. However, despite the models appeal, very few studies have yet tested this hypothesis directly.

Working-memory maintenance. Changes in theta and alpha power have also consistently been linked to memory control processes (Jensen, Gelfand, Kounios, & Lisman, 2002; Ole Jensen & Tesche, 2002; Raghavachari et al., 2006; Sauseng et al., 2010; Wu et al., 2007). For example, Raghavachari and colleagues (2006) recorded from electrode grids at the cortical surface in a group of epileptic patients, while their participants performed a Sternberg working memory task. The recordings indicated a sustained increase in theta power during the entire trial, linking low frequency oscillatory changes directly to working memory processes. Similar findings have been reported in human MEG (Jensen & Tesche, 2002), suggesting that theta oscillations are sensitive to the amount of information in working memory. It should be pointed out though, that some studies also find effects relating to the number of items in memory in the alpha band (Jensen et al., 2002). It is important to realize that these findings are event-related *increases* in alpha power. This is different from the event-related *decrease* that was discussed in the previous section. Jensen and colleagues hypothesize that alpha power contributed to working-memory maintenance by inhibiting irrelevant sensory information (Jensen & Mazaheri, 2010; Jokisch & Jensen, 2007)

More recent studies have demonstrated that theta oscillations might not only be involved in maintaining, but also integrating information in working memory. For example, Wu and colleagues (2007) recorded EEG in participants who were asked to remember a number of locations on the screen (indicated by parenthesis), and a set of letters. Letters were either presented at the same locations a participant needed to remember (bound), or at different locations on the screen (separate). The authors found a larger power increase in the theta band in the bound condition, that is, when location and letters were successfully coupled. Furthermore, there was an increase in coherence between frontal and temporal electrode sites. Based on these findings, the authors conclude that large-scale theta oscillations support binding mechanisms in working memory. The study by Wu and colleagues fits into a larger body of research arguing that

theta oscillations could reflect a general memory control process that is involved in coordinating information from distributed cortical networks (for a review see Sauseng et al., 2010). This interpretation is also in line with research showing a close coupling between theta phase and gamma power during memory operations (Lisman & Idiart, 1995; Nyhus & Curran, 2010; Voytek et al., 2010). Theta oscillations could gate memory encoding and retrieval in local neuronal assemblies through high-frequency gamma oscillations.

In conclusion, oscillatory dynamics in the theta, alpha/beta, and gamma band have been linked repeatedly to memory retrieval and memory control processes. While theta oscillations could be involved in coordinating and possibly integrating distributed cortical networks, alpha/beta and gamma oscillations are particularly sensitive to the stimulus type and thus could be involved in local memory encoding and retrieval processes. The next section will relate some of these dynamics to language processing.

Lexical-semantic retrieval. Memory representation and control are closely intertwined with the ability to understand and produce language. That is, understanding word meaning requires retrieving information from long-term memory, integrating multiple words is required for sentence processing, and following a conversation requires working memory maintenance and control. However, there is currently very little consensus with respect to how oscillations at different frequencies are related to different components of language processing.

At the word level, frequency modulations in the theta, beta and gamma range have been associated with lexical-semantic retrieval (Bastiaansen, Mazaheri, & Jensen, 2008; Bastiaansen, van Berkum, & Hagoort, 2002; Bastiaansen et al., 2005; Bastiaansen, Oostenveld, Jensen, & Hagoort, 2008; Bastiaansen & Hagoort, 2006; Hagoort, Hald, Bastiaansen, & Petersson, 2004; Hald, Bastiaansen, & Hagoort, 2006; Pulvermüller, Lutzenberger, & Preissl, 1999; Van Elk, Van Schie, Zwaan, & Bekkering, 2010; Wang, Zhu, & Bastiaansen, 2012; Wang, Jensen, et al., 2012). For example, Bastiaansen and colleagues (2005) compared event-related oscillatory dynamics in EEG between open and closed class words during story comprehension. The results revealed an increase in theta power and decreases in alpha and beta power in response to open versus closed class words. Only the modulation in the theta band revealed a topographically different effect in open versus closed class words. Based on these results, the authors conclude that theta oscillations most likely reflect lexical-semantic retrieval, while the modulation in alpha could reflect sensory properties of the stimuli. However, since open and closed class words also differ in other ways such as grammatical function, length, word frequency, and neighbourhood size, the interpretation the authors present is not

straightforward. In addition, it is unclear what studying cortical oscillations in this approach contributes that other methods (ERP's, fMRI) could not.

In a different study, Bastiaansen et al. (2008) demonstrate that reading words with strong auditory associations elicits a stronger increase in theta power over temporal sites, while words with strong visual associations show enhanced theta power over occipital electrodes. The authors interpreted these findings as evidence that theta oscillations are selectively involved in the retrieval of sensory-semantic information. In a similar design, Pulvermüller and colleagues (1999) compared nouns with strong visual associations to verbs denoting actions. While nouns showed a relative increase in gamma power as compared to verbs over occipital electrode sites, this pattern was reversed over central sites. Although the design is very similar to Bastiaansen et al. (2008), it should be noted that the conditions in the Pulvermüller study do not only differ in terms of their semantic content but also word category and grammaticality. However, the two studies are not necessarily in conflict with each other. As pointed out in the previous section, there is a tight coupling between theta and gamma oscillations during memory processes and oftentimes a power increase is observed in both bands. Taken together, both Pulvermüller et al. (1999), and Bastiaansen et al. (2008) provide evidence for topographic differences, as a function of modality-specific word content. However, as neither of the two studies attempted to localize the effect in source space, these results should be considered with caution.

A somewhat different spectral profile has been reported for action verbs. For example, in a study by van Elk et al. (2010), participants were presented with sentences including an action verb in an animal or a human context. The premise of the study was that there is a difference in how familiar the participant is with the action in either of the two contexts. Contrasting these two conditions, the authors observed stronger desynchronization in alpha and beta bands in the animal as compared to the human context. Critically, similar spectral changes have been reported during action planning and execution (Hari et al., 1998; Neuper et al., 2006; Pfurtscheller and Neuper, 2010). The authors interpret these results as evidence that motor processes support lexical-semantic retrieval and integration. However, there are two fundamental problems with this study. First, human participants are much more familiar with the actions of other human beings, therefore motor related spectral components should be stronger in the human context as compared to the animal context. The fact that the authors found the reversed pattern of results makes the study difficult to interpret in this framework. Second, the authors find a strong correlation between motor related spectral modulations and cloze probability. This suggests that the difference is more likely to be the result of sentence level processes rather than lexical semantics.

Sentence-level integration. Understanding the meaning of words requires the ability to retrieve conceptual knowledge from long-term memory. However, the demands of everyday language understanding go beyond retrieving isolated chunks of information. In order to understand a sentence, or even a story, the language user needs to maintain, integrate, and respond to this information in the blink of an eye. Accordingly, oscillatory dynamics during sentence comprehension are thought to reflect these control and integrative mechanisms.

A common method in the electrophysiological investigation of language functions is the comparison of control sentences with sentences that contain a grammatical (Osterhout & Holcomb, 1993) or semantic anomaly (Kutas & Hillyard, 1980). A range of studies have used this paradigm to show that semantic violations elicit enhanced oscillatory activity in the theta and gamma band, whereas decreases in the beta band are more common during grammatical violations. For example, Hagoort and colleagues (2004) compared oscillatory responses to three different types of sentence types. The sentences were exactly the same except for the last word, which could be a semantic violation (Dutch trains are *sour*), a world knowledge violation (Dutch trains are *white* – Dutch trains are yellow, and every participant knew this), or no violation (Dutch trains are *yellow*). The authors found that semantic violations elicit a stronger increase in theta power while world knowledge violations induce enhanced gamma activity, suggesting that different processes could be involved in retrieving the different types of information. In a more recent study, Wang and colleagues (Wang, Zhu, et al., 2012) used a similar paradigm in which the final word of the sentence was predictable, unpredictable, or incongruent. The authors show that theta oscillations are more sensitive to semantic violations, while gamma oscillations are reduced if the upcoming word is unpredictable. In line with previous sections (Engel et al., 2001; Tallon-Baudry & Bertrand, 1999), one could argue that gamma oscillations in both tasks operate at the level of perceptual-semantic information retrieval. For example, a highly predictable sentence might elicit a perceptual model of the sentence meaning more readily than a sentence that is unpredictable (Zwaan, 2003). Theta oscillations in turn might be involved in the accumulation of semantic knowledge more generally. Additional support for this claim is the discovery that only theta power shows a linear increase over the course of sentence processing (Bastiaansen, van Berkum, & Hagoort, 2002)

While enhanced theta and gamma power are often observed during semantic violations, reduced beta and alpha power are more frequently reported during grammatical violations. For example, Davidson and Indefrey (2007) presented participants compared control sentences to sentences with either a semantic or a grammatical violation. In line with previous findings, semantic violations elicited a

power increase in the theta band. In contrast, grammatical violations elicited a decrease in the alpha and beta band. These results are in line with other studies showing effects in the beta band to grammatical violations (Bastiaansen, Magyari, & Hagoort, 2010). A different way to study grammatical processing in the brain is to compare sentences of different complexity. For example, Weiss and colleagues (Weiss et al., 2005) have compared oscillatory responses to sentences of different complexity (center-embedded versus relative clauses), and demonstrated that more complex sentences (center-embedded) show a change in connectivity between anterior and posterior electrodes covering a wide range of frequencies (theta, gamma, and beta). Thus, although research on grammatical processing has mainly found evidence for a contribution of beta oscillations, these results are even less conclusive than findings at the word level.

The previous section has introduced the *temporal binding hypothesis* (Singer and Gray, 1995), a framework which explains how synchronized neuronal activity could benefit perceptual feature integration. In addition, evidence was presented that such a mechanism in humans is reflected in oscillatory dynamics in the gamma range (Tallon-Baudry and Bertrand, 1999). The current section has reviewed evidence suggesting that similar mechanisms may also be important for memory retrieval, working memory, and language processing. For example, local changes in high frequency gamma oscillations in visual areas have been associated with successful retrieval of perceptual information (Osipova et al., 2006). In addition, a number of studies have related memory retrieval, and integration with slow cortical oscillations in the theta band (Wu et al., 2007; Sauseng et al., 2010). It has been argued that these slow oscillations could be involved in coordinating information at a larger spatial scale, that is, across distributed cortical networks. The idea that low and high frequency oscillations operate at different spatial scales is not new (von Stein and Sarnthein, 2000; Donner and Siegel, 2011), but so far few studies have used this framework to study language understanding through cortical oscillations.

Although previous studies have reported oscillatory changes at multiple frequencies for language-related phenomena (e.g., semantic/grammatical violations, processing of open-class words), it is often not clear why a change in a particular frequency band would be expected. In addition, the reasons for using oscillatory dynamics to study language processing are often poorly motivated, which might have led to the variability in the findings. Specifically, cortical oscillations in a specific frequency band may support multiple cognitive processes, while a single process likely induces changes in multiple frequency bands. Given this many-to-many mapping between oscillatory dynamics and cognitive processes it becomes even more important to formulate hypotheses in terms of the underlying network dynamics, rather than

linking individual frequency-bands to cognitive phenomena directly. Examples of the latter are studies who have linked cortical oscillations at specific frequency bands to semantic violations (Wang et al., 2012a, 2012b), or even a particular word class such as action words (Van Elk et al., 2010; Moreno et al., 2013).

However, as the memory literature shows, cortical oscillations can reveal unique and novel insights, provided the hypotheses are formulated in terms of the underlying neurophysiology (Axmacher, Mormann, Fernández, Elger, & Fell, 2006; Hanslmayr et al., 2012; Jensen & Mazaheri, 2010). For example, the idea that low and high frequency oscillations operate at different scales has profound implications for the problem of how distributed semantic information is integrated into coherent conceptual representations. Particularly, the level at which information is integrated or processed (local modality-specific, or global supramodal), should be reflected in oscillatory dynamics at different frequencies. The last section will provide an outline of the empirical work presented in the present thesis and how oscillatory dynamics are used to study conceptual integration in the brain.

Thesis outline and aims

The first section of this review has presented evidence that word meaning is represented in a distributed cortical network, which may be supported by a central hub, or multiple convergence zones. Yet, the mechanism of how distributed word content is integrated into a coherent concept is still poorly understood. The second section has discussed evidence that synchronized neuronal activity, which is reflected in oscillatory neuronal activity at the level of the scalp, could be involved in integrated perceptual information as well as information that is stored in long-term memory. But, while oscillatory dynamics at different frequencies have been frequently linked to language understanding, there is currently no coherent mechanistic framework of how oscillatory neuronal activity contributes to word retrieval. The aim of the current thesis is to explore how cortical oscillations at different frequencies contribute to the integration of word meaning at different cortical scales. Furthermore, the current thesis aims to provide a mechanistic and physiologically plausible model of semantic feature integration.

Chapter 2 will address how the complexity of perceptual information, encoded in word meaning, is reflected in oscillatory neuronal activity in EEG. According to the information via desynchronization framework (Hanslmayr et al., 2012), the amount of information that is currently represented in the system should be reflected in reduced oscillatory activity in the alpha/beta band. This hypothesis is tested by asking

participants to generate perceptual features from different modalities in response to a concept cue.

Chapter 3 directly addresses whether integrating information across multiple semantic networks is reflected in slow cortical oscillations in the theta band. In a dual property verification task, participants are presented with two features, followed by a target word. Critically, the features are either from the same or two different modalities. In two experiments, reaction times and EEG are recorded. It is hypothesized that information that is more widely distributed in the brain will take longer to be processed, which should elicit longer reaction times in the multimodal condition. Additionally, integrating two features from multiple modalities should engage low-frequency oscillatory dynamics that are involved in gating or control mechanisms during memory operations.

Chapter 4 builds on the results from Chapter 3 by investigating whether oscillatory dynamics during multimodal conceptual integration reflect semantic integration or conceptual switching. Dual property verification is used to test whether the modality-specificity effect is specific to feature pairs that are congruent with a target word (integration), or whether the effect is observed independent of feature congruency (switching).

Chapter 5 presents a further extension of previous chapters using MEG. MEG offers a much better spatial resolution in source space as well as a sufficient signal-to-noise ratio in higher frequency bands. Specifically, the aims are to test which areas generate low frequency oscillatory activity during cross-modal semantic feature integration, and whether integrating features from the same modality is reflected in high rather than low frequency bands. In line with previous memory research, this effect should be observed in modality-specific memory networks.

Chapter 6 investigates directly how modality-specific and central hub like structures in the brain communicate when participants retrieve semantic features from the same or different modalities.

Chapter 7 integrates the results from empirical work presented in the current thesis into a neurophysiologically plausible framework. The framework describes how oscillatory dynamics at different frequencies contribute to the selection and integration of semantic features from long-term memory.

Chapter 2:

When less is more: Perceptual complexity of lexical-semantic content is reflected in reduced oscillatory neuronal activity

Abstract

In recent years, a plethora of studies have demonstrated that retrieving perceptually complex word meaning engages distributed cortical networks in a modality-specific way. While much is known about *where* word meaning is stored in the brain, the neural mechanisms underlying *how* this information is retrieved are still poorly understood. The current study contributes to a better understanding of this problem by investigating how differences in perceptual complexity between modality-specific semantic content is reflected in dynamic changes in oscillatory neuronal activity. Participants were presented with concrete nouns and asked to generate two semantic features from one of three modalities (visual, auditory, haptic). The informational complexity (entropy) in the features that participants used to describe an object in a given modality was taken as a proxy for the perceptual complexity of each modality. The EEG signal was recorded and used to compute event-related potentials, and changes in oscillatory neuronal activity while participants retrieved modality-specific features. The results indicate that retrieving perceptually complex semantic content is reflected in suppressed oscillatory neuronal activity in the beta band (16-30Hz), as well as an early modulation (220-240ms) in event-related potentials. These results are in line with a recently proposed account arguing that suppressed oscillatory power or de-synchronization in the alpha and beta band is associated with richer informational content in memory.

Introduction

Semantic memory comprises an individual's knowledge about the objects and entities in the world. Functional imaging studies over the last years have shown that semantic knowledge is at least partially organized in modality-specific cortical networks (Martin and Chao, 2001; Patterson et al., 2007; Pulvermüller and Fadiga, 2010; Binder and Desai, 2011). However, while much is known about *where* semantic information is stored, the neuronal dynamics supporting *how* this information is retrieved in a dynamic and flexible way are still poorly understood. For example, modalities differ in perceptual complexity, and this is reflected in the relative number of words we have in our mental lexicon to describe e.g. what an object looks like (visual) versus what it feels like (haptic) (see relative feature quantities in Lynott & Connell, 2009; van Dantzig, Cowell, Zeelenberg, & Pecher, 2010). Investigating how perceptual complexity affects the neuronal dynamics of the brain during word processing is a vital step in understanding the format of lexical-semantic representations and how it affects lexical retrieval.

Numerous imaging studies in previous years have shown that accessing word meaning engages modality-specific networks in the brain. For example, action words have been shown to engage a fronto-parietal motor system (Hauk et al., 2004; Postle et al., 2008; Rüschemeyer, Brass, & Friederici, 2007), color words activate parts of the ventral visual system (Simmons et al., 2007), and words with acoustic associations activate auditory areas in superior and middle temporal lobes (Kiefer et al., 2008). However, while this is compelling evidence for the existence of a distributed semantic network, these findings do not explain *how* information is represented in different modalities.

As many embodied theories of language postulate (Barsalou, 2008; Pulvermüller and Fadiga, 2010), one possibility is that word meaning is stored in a perceptual format. It follows that modality-specific networks should be sensitive to the complexity with which perceptual information is encoded in word meaning. This has been demonstrated for visual and action related information (Gauthier et al., 1997; van Dam et al., 2010). For example, Gauthier and colleagues reported that nouns denoting less richly encoded basic level categories show less activation in visual cortices than nouns denoting more specific subordinate categories. Following up on this idea, van Dam and colleagues showed a similar effect for subordinate versus superordinate action verbs in the inferior parietal lobule, which is part of a fronto-parietal motor network. These results suggest that retrieving modality-specific semantic content is sensitive to the perceptual complexity of a word.

However, the impetus to study modality-specific word meaning has led to an artificial emphasis on stable semantic concepts from one dominant modality. In contrast, more recent studies suggest that the recruitment of modality-specific networks is highly dependent on the context (Hoenig et al., 2008; Raposo et al., 2009; van Ackeren et al., 2012; van Dam et al., 2012). For example, van Ackeren and colleagues showed that non action related sentences such as *It is hot in here* activate action-related semantic features if presented as an indirect request to open a window (van Ackeren et al., 2012). In addition, van Dam and colleagues showed that the relative activation in motor networks in response to a target word is different when participants think about the function of an object versus its color (van Dam et al., 2012). But modalities also differ in terms of their perceptual complexity. For example, the resolution of the human visual system is far superior to the haptic, or olfactory system, and this is also reflected in the number of words/features we can use to describe what an object looks (visual) like versus what it feels like (haptic). Thus, an attention shift from a perceptually poor modality to a perceptually rich modality also requires a shift in perceptual complexity, and the amount of retrieved features.

One way to conceptualize perceptual complexity physiologically is in terms of the amount of information that is encoded in the underlying neuronal populations. According to the *information via de-synchronization hypothesis* (Hanslmayr et al., 2012), the amount of information that is encoded or retrieved from memory is inversely related to spectral power in the alpha (10-14 Hz) and beta band (16-30 Hz). It is widely accepted that an increase in spectral power reflects enhanced neural synchronization in the underlying populations. However, if multiple neurons fire in synchrony, there is an overall reduction in the amount of information, or entropy (Shannon and Weaver, 1949), which is encoded in the signal. Conversely, a decrease in power, or neuronal de-synchronization, should reflect an increase in information. Indeed, empirical evidence suggests that retrieving items that were initially encoded with more information is associated with a stronger reduction in spectral power than items that were encoded with less information (Khader & Rösler, 2011; Waldhauser, Johansson, & Hanslmayr, 2012). An interesting prediction can be derived from this idea: Accessing information from perceptually complex modalities (e.g., vision) will be reflected in more suppressed oscillatory power than accessing information from perceptually less complex modalities (e.g., touch).

The primary aim of the current study is to test whether retrieving information from perceptually complex modalities is reflected in suppressed oscillatory neuronal dynamics in the alpha/beta band. EEG is used to measure oscillatory neuronal activity while participants retrieve two visual, auditory, or haptic features of a given target word.

After the experiment, a measure of informational complexity (entropy) is computed from the participants' responses for each modality. In line with the information via de-synchronization hypothesis (Hanslmayr et al., 2012), we predict that modalities with perceptually complex feature content show a stronger suppression in oscillatory alpha/beta power. A secondary goal of the study was to explore *when* perceptual complexity effects arise during language processing. To this end event-related potentials were compared when participants think about features of an object from different modalities.

Methods

Participants

Participants were sixteen healthy individuals (5 male) with no known neurological disorder and normal or corrected to normal vision. The age range was between 18 and 27 years ($M=20.18$). All participants were students at the University of York, and participated on a voluntary basis. Participants received either a financial reward or course credits for taking part in the experiment. All participants gave written informed consent according to the Declaration of Helsinki and were debriefed after the study. The study was approved by the Ethics Committee of the York Neuroimaging Centre.

Experimental design

Participants performed a feature generation task. After a brief baseline period (1000 ms) the participant saw a visual cue (visual, auditory, haptic), followed by a delay (500 ms) and a target noun (1500 ms). The participants were instructed to think about features of the target word from the cued modality. After each trial, the participant was prompted to enter two modality-specific features using the keyboard (Figure 2.1). The experiment was preceded by 10 practice trials to familiarize participants with the experiment.

The experimental items consisted of 180 nouns referring to different semantic categories (100 artificial objects, 40 animals, 26 natural objects, 8 types of food, and 6 body parts). As each noun was presented in all of the three conditions (Visual, Auditory, Haptic), no further matching of the words was required. In order to prevent priming effects, each participant saw a target word only with one of the three cuing conditions. The order of nouns and conditions was pseudo-randomized within participants and the pairings between nouns and conditions was counterbalanced across participants.

Data recording and pre-processing

EEG was acquired from 64 Ag-AgCl electrodes that were positioned on an electrode cap according to a 10-20 system. All electrodes were re-referenced offline to the algebraic average of the two mastoids. Horizontal and Vertical eye movements were recorded with a set of bipolar Ag-AgCl electrodes. The signal was amplified using an ANT amplifier. Impedances of the cortical electrodes were kept below 10 k Ω . The signal was recorded with a sampling frequency of 500 Hz.

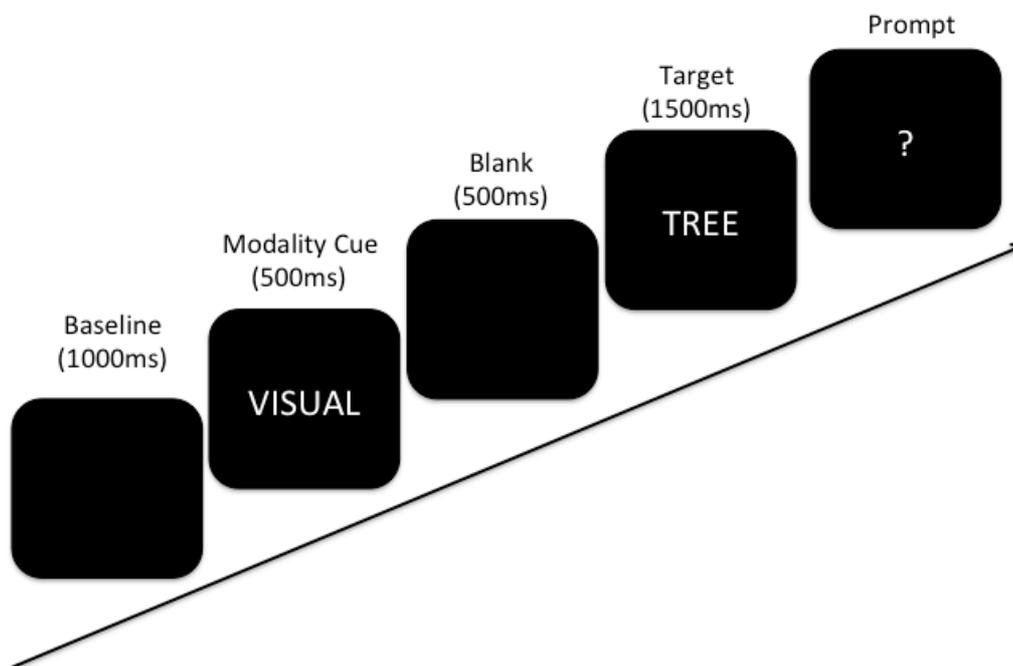


Figure 2.1. Experimental design. The trial started with a period of rest (1000 ms), which served as baseline for subsequent spectral analysis. The baseline period was followed by a modality-cue (Visual, Auditory, Haptic) (500 ms), a delay period (500 ms), and the target noun (1500 ms). After each trial, participants were asked to report two features of the target word from the cued modality.

Offline analyses were conducted using Matlab 7.14 (Mathworks, Natick, MA) and Fieldtrip, a Matlab toolbox for analyzing EEG/MEG data (Oostenveld et al., 2011). For subsequent analysis, the data were band-pass filtered (0.05-170 Hz, Butterworth filter, low-pass filter order = 4, high-pass filter order = 3). Line-noise was suppressed by filtering the 50, 100, and 150 Hz Fourier components. Artifact rejection was performed in three consecutive steps. First, muscle artifacts were removed using semi-automatic artifact rejection. Subsequently, extended infomax independent component analysis

(ICA), with a weight change stop criterion of $< 10^{-7}$, was performed to identify, and reject ocular components. One participant was excluded from the analysis due to excessive muscular and ocular artifacts. In addition, seven trials had to be excluded from each dataset due to a technical error.

Analysis of event-related potentials

For the analysis of event-related potentials, the data were further low-pass filtered at 30 Hz. Subsequently; data were segmented into epochs from 150 ms before stimulus onset to 600 ms after target onset and corrected with a pre-stimulus baseline of 150 ms.

Statistical analysis of event-related potentials was performed using a two-step procedure. First, in order to objectively identify evoked components for subsequent analyses, the global root mean square (RMS) was computed. Data from all conditions were averaged, and the square root was computed for the mean of the squared signal amplitude across electrodes at each point in time. Local peaks were identified at ~110 ms, ~150 ms, 230 ms, and ~350 ms, which fall in to the time windows of the N1, P2 and N4 (Figure 2.2). Given a priori knowledge about the morphology of these peaks, 20 time windows of interest (TOI) were chosen for the first three components, and a 100ms TOI for the last component. Second, separate analyses of variance with repeated measures (ANOVA) were computed for each electrode position and subjected to a cluster-randomization technique (Maris and Oostenveld, 2007). Cluster-randomization effectively reduces the number of comparisons by combining neighboring electrodes above a given threshold ($\alpha = .05$). The summed F-values are compared to a permutation distribution in which the condition labels are permuted on 1000 iterations. Subsequently a maximum statistic is applied to correct for the number of clusters. This procedure is highly objective, as it does not require the manual selection of electrodes for statistical comparisons. Lastly, to investigate the possible effect of feature entropy for each of the three conditions, linear regression was applied to the time windows of interest and subjected to a cluster-based permutation analysis.

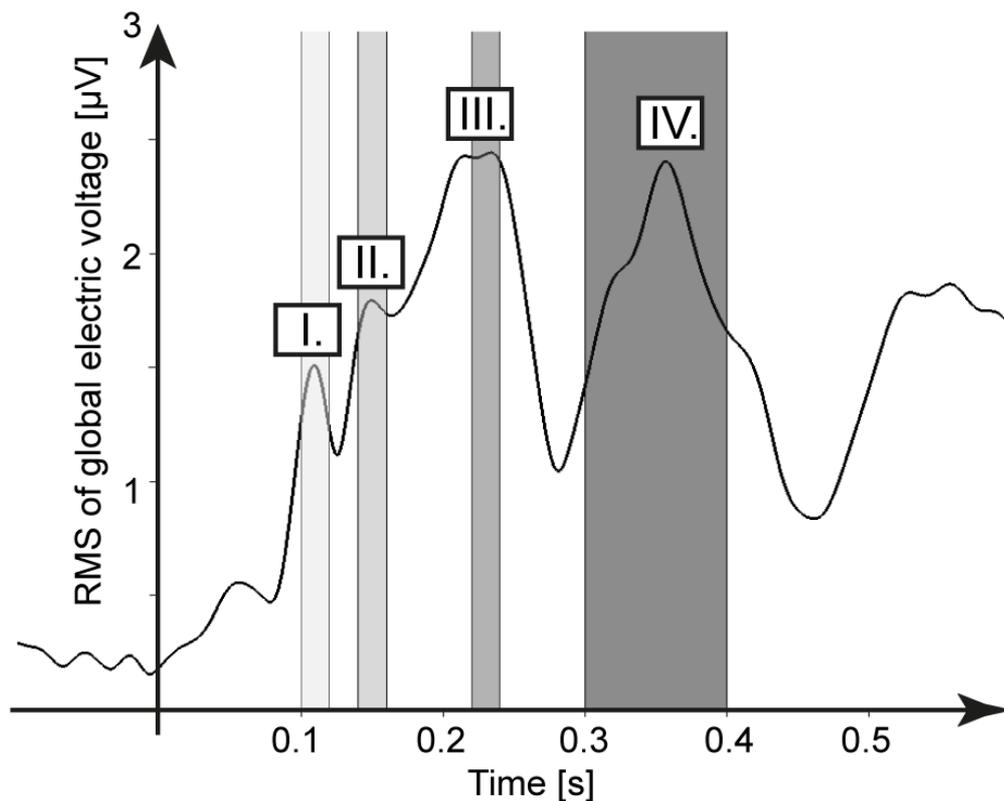


Figure 2.2. Global root mean square amplitude. RMS was calculated across all conditions and electrodes. Local peaks were found at (I) ~110ms, (II) ~150ms, (III) ~230ms, and (IV) ~350ms. 20ms time windows of interest was chosen for the first three peaks (I, II, III), and a 100ms time window for the last peak (IV) (shaded bars).

Analysis of spectral power

Dynamic changes in spectral power during the presentation of the target word were computed by convolving the signal with a complex Morlet wavelet (5 cycle width) using 50ms time steps and a step size of 2Hz in the frequency domain. These transformations were performed at the individual trial level and reflect both evoked and induced components of the signal. Subsequently, trials were averaged for each condition and subject, and percentage signal change was computed using a time window between 500 and 200 ms before the onset of the cue. The baseline normalization procedure is equivalent to the event-related de-synchronization technique (Pfurtscheller & Lopes da Silva, 1999), except that positive values denote synchronization, and negative values de-synchronization.

Similar to the analysis of event-related potentials, ANOVA with repeated measures was used in combination with a cluster randomization technique to identify

spectral power changes that are tuned to memory search in a given modality. Cluster-based permutation was performed on 2 frequencies of interest (alpha: 8-14 Hz, and beta: 16-30 Hz) in a time window from 200-400ms after stimulus onset. Clustering was performed across space (electrode locations)

Results

Behavioral analysis

Each participant supplied two features for a given target word from one of three cued modalities (visual, auditory, haptic). For subsequent analyses, these responses were cleaned in the following way. Any spelling errors or inconsistencies in spelling were corrected. Further, words with different morphological inflections were converted into a single form (e.g., *jingle*, *jingly*, *jingling* all became *jingling*). Any onomatopoeia (e.g., *ding ding*, *broom broom*) were removed from the data if they did not resemble an existing lexical item in the English language. Lastly, responses such as *car sound* (for auditory feature of car), or *car shaped* (for visual feature of car) were excluded from any further analyses.

Subsequently, the entropy of these features was computed for each condition (visual, auditory, haptic), and participant. Entropy (H) is a mathematical concept from information theory that quantifies the unpredictability or complexity of informational content as:

$$H(X) = -\sum P(x_i) \log_2 P(x_i)$$

where $P(x_i)$ is the probability of a specific event in a set of events (X). Here $P(x_i)$ is the probability that a participant uses a given feature in one of three modalities. Concretely, if a participant uses few features, many times, there is less information in the responses (low entropy) than if a participant uses many different features only a few times (high entropy).

To test whether different semantic modalities are associated with different levels of informational complexity, feature entropy was subjected to analysis of variance (ANOVA) with repeated measures. The test statistic revealed a main effect of cued modality on semantic entropy ($F(2,28)=60.21$, $p<.001$ partial $\eta^2=.513$). To further explore this effect, post-hoc tests were applied using FDR-adjustment to correct for multiple comparisons. Feature entropy was higher in both visual and auditory conditions than in haptic conditions (both $p < .001$), yet there was no difference between the visual

and auditory condition ($p = .13$). These results confirm the notion that the visual and auditory modalities encode perceptually rich information, while the haptic modality encodes perceptually poorer information in comparison.

Table 2.1. Behavioral results depicted as the mean entropy of semantic features in each of the three priming conditions.

Modality Cue		Feature entropy
Visual	M	5.91
	SE	0.09
Auditory	M	5.80
	SE	0.08
Haptic	M	4.94
	SE	0.16

Event-related potentials

100-120ms. There was a negative deflection in the event-related time course around 110ms. However, no clusters were identified in this time window by our clustering algorithm. In addition, no significant clusters were found for the feature entropy regression in any of the three conditions.

140-160ms. Despite a minor divergence between the visual condition and the auditory and haptic conditions around 150ms, the clustering algorithm did not identify a cluster in this time window.

220-240ms. There was a positive peak in the time window around 230 ms. The cluster-based randomization procedure identified one significant cluster ($p=.037$, cluster-corrected; Figure 2.3), which was centered over left posterior electrodes, suggesting that memory search is reliably constrained by the modality-specific cue in an early time window corresponding to the P2 event-related component. Post hoc tests revealed that there was a stronger deflection for the visual and auditory as compared to haptic cues ($p=.01$, and $p=.024$, respectively), while there was no difference between visual and auditory cues ($p=.846$). All p-values are FDR-adjusted.

300-400ms. The last time window around 350ms revealed a negative deflection in the signal. Cluster-based permutation revealed one significant cluster over central and

posterior electrodes ($p=.036$, cluster-corrected; Figure 2.3), suggesting that the effect of modality-constrained memory search leaks into the time window corresponding to the N400 effect, which has been associated consistently with semantic processing. Post hoc tests revealed that there was a stronger negative deflection for the visual and auditory as compared to haptic cues ($p=.01$, and $p=.03$, respectively), while there was no difference between visual and auditory cues ($p=.865$). All p -values are FDR-adjusted.

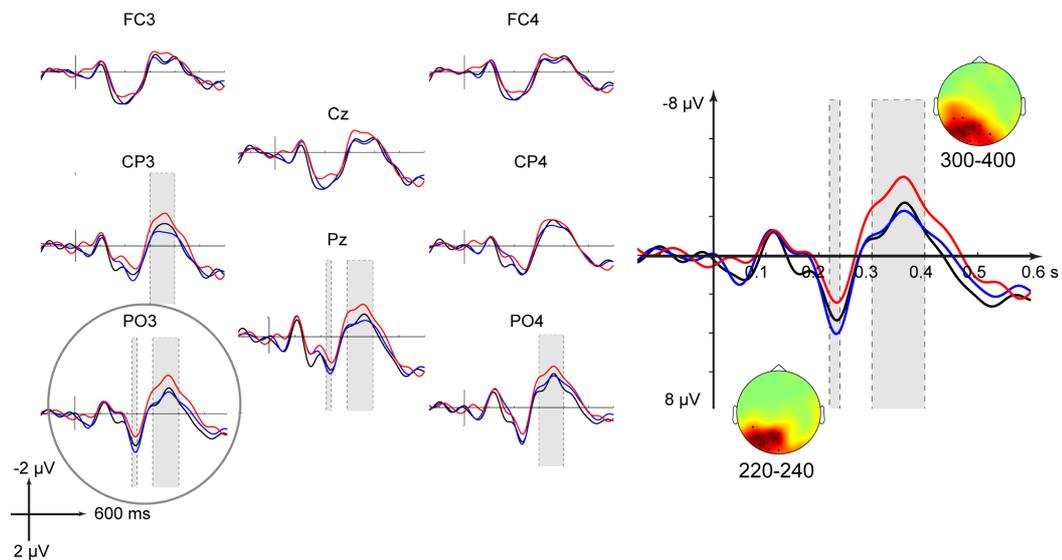


Figure 2.3. Event-related potentials during target presentation. The left panel shows the event-related time course for 8 representative electrodes during the presentation of the target in each of the three cue conditions (Visual: black; Auditory: blue; Haptic: red). The left panel depicts a magnified illustration of the event-related time course at electrode PO3. Shaded areas depict the early (220-240ms) and late (300-400ms) time window in which the clustering algorithm identified a significant interaction between conditions. The distribution of each effect is represented in the adjacent topographic scalp maps. Black dots depict channels that are part of the significant cluster (mean time courses were smoothed at 10Hz for illustration purposes).

Spectral analysis

ANOVA with repeated measures in combination with cluster-based permutation was performed in a time window from 200 to 400 ms after the presentation of the target word in two frequency bands of interest. The clustering algorithm identified one significant cluster in the beta band (16-30 Hz) ($p = .028$, cluster-corrected) for the effect of modality-cueing. Post-hoc tests revealed that there was a significantly stronger beta

suppression for the visual and auditory condition versus the haptic condition ($p=.015$, and $p=.019$, respectively), while there was no difference between the visual and auditory condition ($p=.61$).

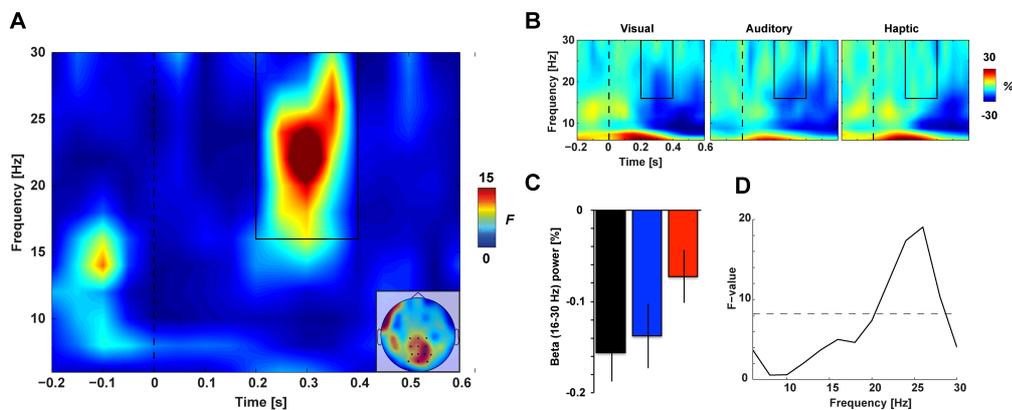


Figure 2.4. Spectral decomposition showing the effect of the cued modality. A. Time-frequency decomposition showing the effect of the cued modality (rectangle) on the target word averaged over all electrode locations in the cluster. The topographical representation at the right bottom depicts the spatial extend of the cluster (black dots). B. Time frequency representations for each of the three conditions individually. During the time period of interest (rectangle) there is a reduction in all conditions relative to baseline. C. Bar graphs show the average signal decrease in the beta band for all three conditions. The graph shows that the visual (black) and auditory (blue) conditions show a larger decrease in beta power as compared to the haptic condition (red). D The spectrogram shows the distribution of the effect of cued modality in the frequency domain during the time period of interest. The illustration shows that the effect reaches significant (depicted as dashed line) between 20 and 30 Hz with a peak at 26 Hz.

Discussion

The aim of the current study was to test whether accessing lexical-semantic features from perceptually complex modalities is reflected in suppressed spectral power in the alpha/beta band. To this end, EEG was recorded while participants were asked to retrieve two perceptual features from one of three modalities (visual, auditory, haptic) for a given target word. Perceptual complexity was estimated as the entropy in the features that participants generated in a given modality. Entropy was similar in the visual and auditory modality, but reduced in the haptic modality. This pattern was reflected in more suppressed oscillatory power in the beta range (16-30 Hz) for perceptually rich modalities (visual, auditory) compared to the perceptually poor modality (haptic). Analyses of event-related potentials revealed a similar pattern around

~230, and ~350 ms, suggesting that context-dependent perceptual features are retrieved early during lexical-semantic processing. The implications of these findings are discussed below.

Oscillatory power diminishes as information increases

Reduced oscillatory neuronal activity in the alpha/beta band (~8-30 Hz) has been reported in a range of cognitive tasks, including language processing (Davidson and Indefrey, 2007; Van Elk et al., 2010; Wang et al., 2012a; Moreno et al., 2013). However, the cognitive function of this phenomenon is still poorly understood. According to a recent account, oscillatory de-synchronization reflects the amount of information that is encoded/retrieved at a given moment in time (Hanslmayr et al., 2012). In line with this hypothesis, the current study reports a reduction in spectral power in the beta band (16-30 Hz) when participants think about lexical-semantic features from perceptually complex (e.g., visual, and auditory features) modalities. In other words, the pattern of de-synchronization reflects the entropy of the features that participants use to describe visual, auditory, or haptic aspects of an object. While beta de-synchronization has been shown repeatedly to reflect encoding and retrieval complexity in memory research (Hanslmayr et al., 2012; Waldhauser et al., 2012), the present study is the first to directly apply these concepts to the retrieval of lexical-semantic content.

As multiple studies have reported beta suppression during action observation and execution (Hari et al., 1998; Neuper et al., 2006; Pfurtscheller and Neuper, 2010), research on lexical-semantics has focused primarily on the role of beta oscillations in action word processing. For example, Moreno and colleagues found reliable alpha and beta suppression when participants listened to action-related sentences versus abstract sentences, and this effect resembled the pattern of de-synchronization during action observation (Moreno et al., 2013). Similar to functional imaging studies, alpha and beta de-synchronization have also been shown to be sensitive to the context in which action verbs are presented (Van Elk et al., 2010). These findings are not in conflict with the current study. For example, in the study by Moreno and colleagues (2013), action sentences clearly contained more complex perceptual content than abstract sentences. The current study extends this research showing that beta suppression is indicative of perceptual complexity more generally - that is perceptual complexity in multiple modalities. This is in line with a growing body of evidence suggesting that beta oscillations are involved in multiple cognitive processes, without motor involvement (Engel and Fries, 2010; Donner and Siegel, 2011; Hanslmayr et al., 2012).

A different line of research has shown suppression in beta power during grammatical and semantic expectation violations (Davidson and Indefrey, 2007; Wang

et al., 2012a). Violations require a form of reanalysis or adaptation, which involves additional information retrieval. This interpretation would be in the spirit of the present study. However, expectancy violations in language induce a complex change in the frequency domain covering multiple frequency bands, and other studies have emphasized the role of different frequency bands, such as theta, in expectancy violations (Hald et al., 2006; Wang et al., 2012b). For these reasons, a direct comparison to the current study remains difficult.

In sum, in line with the information via de-synchronization hypothesis (Hanslmayr et al., 2012), and previous memory research (Waldhauser et al., 2012), the present findings show that retrieving information from perceptually rich modalities during word processing induces enhanced beta suppression. While previous research has reported beta suppression during action verb processing, the current study extends this work suggesting that beta suppression might be a more general mechanism that is linked to the complexity of information retrieval.

Complex perceptual information is activated early during language processing.

Early electrophysiological research using the event-related potential technique has reported effects of semantic task manipulations around 400ms after stimulus onset (Kutas and Hillyard, 1980; Kutas and Federmeier, 2000). Yet, more recent evidence suggests, that retrieving modality-specific word content may already start around 150 ms (Pulvermüller et al., 2005b, 2009; Kiefer et al., 2008; Pulvermüller and Shtyrov, 2009). For example, Kiefer and colleagues found that words with strong auditory associations engage left superior temporal areas as early as 150ms (Kiefer et al., 2008). Further, Pulvermüller and colleagues have shown that motor areas in the cortex are sensitive to lexical-semantic retrieval of words with action-specific content as early as 150ms (Pulvermüller et al., 2005b; Pulvermüller and Shtyrov, 2009).. The present study is in line with these findings showing an early effect of perceptual complexity around ~230ms.

Relation to semantic richness

In the current study, perceptual complexity was operationalized as the entropy in the feature words that participants used to describe modality-specific aspects of the target word. The concept of perceptual complexity in the current study relates to earlier work on semantic richness. While there are multiple metrics for semantic richness, the most commonly used is the number of features that are associated with a particular word.

Semantic richness has been shown to modulate both naming and lexical decision latencies (Pexman et al., 2002, 2003), with faster responses to words with richer

semantic content. More recent studies have emphasized that for concrete objects, particularly the number of *perceptual* features seem to modulate the effect, which, in line with the current study emphasizes the role of modality-specific information in lexical-semantic retrieval.

In line with the current study, semantic richness has been shown to modulate lexical-semantic retrieval early (Kounios et al., 2009; Rabovsky et al., 2012). For example, Kounios and colleagues (2009) used the event-related potential technique to show that words with more semantic features, diverge from words with few semantic features already around ~200ms. In the current study a difference between perceptually rich and poor modalities was found as early as 230ms.

However, there are at least two important differences between the current study and previous work on semantic richness. First, semantic richness is quantified as a property of a concept. In contrast, perceptual complexity in the present study captures the amount of information (entropy) that is used to describe concepts by each individual. The term *perceptual* here refers to the use of perceptual modality-specific features. In addition, it should be mentioned that perceptual complexity in the current task is highly context-dependent. That is, how many features are retrieved depends on which aspects of the concept (e.g., which modality) a participant is currently thinking about. Second, while perceptual complexity is often quantified across a population of raters, here perceptual complexity is quantified as the number of features a participant uses for a given target modality, and their probability of use. This is relevant as features may be used equally often, or one feature may be used far more than the others. In sum, while semantic richness and perceptual complexity certainly capture similar aspects of lexical-semantic retrieval, the latter specifically takes into account the dynamic nature and modality-specific organization of the semantic system.

Conclusion

The aims of the current study were twofold. The first aim was to test whether retrieving perceptually rich word meaning is reflected in suppressed oscillatory activity in the alpha/beta band. In line with a recently proposed framework, the information via de-synchronization hypothesis, beta power was reduced for the perceptually more complex visual and auditory modalities as compared to the less perceptually rich haptic modality, suggesting that beta suppression is sensitive to the amount of information retrieved at a more general level. Here, we extend previous research on oscillatory suppression during lexical-semantic retrieval, which has almost exclusively focused on action word/sentence processing. The second aim was to track when in time perceptually rich

semantic features are retrieved. In line with recent evidence (Pulvermüller et al., 2005b, 2009; Kiefer et al., 2008; Kounios et al., 2009; Pulvermüller and Shtyrov, 2009; Rabovsky et al., 2012), we found an early effect (~230ms) of perceptual complexity suggesting that multiple features are activated in a rapid and possibly parallel way.

Chapter 3:

Cross-modal integration of lexical-semantic features during word processing: Evidence from oscillatory dynamics during EEG

(based on: **van Ackeren MJ**, Rueschemeyer S-A (2014) Cross-modal integration of lexical-semantic features during word processing: Evidence from oscillatory dynamics during EEG. PLoS ONE 9(7): e101042. doi: 10.1371/journal.pone.0101042)²

Abstract

In recent years, numerous studies have provided converging evidence that word meaning is partially stored in modality-specific cortical networks. However, little is known about the mechanisms supporting the integration of this distributed semantic content into coherent conceptual representations. In the current study we aimed to address this issue by using EEG to look at the spatial and temporal dynamics of feature integration during word comprehension. Specifically, participants were presented with two modality specific features (i.e., visual or auditory features such as *silver* and *loud*) and asked to verify whether these two features were compatible with a subsequently presented target word (e.g., *WHISTLE*). Each pair of features described properties from either the same modality (e.g., *silver*, *tiny* = visual features) or different modalities (e.g., *silver*, *loud* = auditory, visual). Behavioral and EEG data were collected. The results show that verifying features that are putatively represented in the same modality-specific network is faster than verifying features across modalities. At the neural level, integrating features across modalities induces sustained oscillatory activity around the theta range (4-6 Hz) in left anterior temporal lobe (ATL), a putative hub for integrating distributed semantic content. In addition, enhanced long-range network interactions in the theta range were seen between left ATL and a widespread cortical network. These results suggest that oscillatory dynamics in the theta range could be involved in integrating multimodal semantic content by creating transient functional networks linking distributed modality-specific networks and multimodal semantic hubs such as left ATL.

² The author designed the experiment, analyzed the results, and wrote the article under supervision of Dr. Shirley-Ann Rueschemeyer

Introduction

The embodied framework of language suggests that lexical-semantic knowledge (i.e., word meaning) is stored in part in modality-specific networks that are distributed across the cortex (Barsalou, 2008; Vigliocco et al., 2009; Pulvermüller and Fadiga, 2010; Binder and Desai, 2011). For example, words denoting colors (e.g., *red*, *green*) have been shown to engage parts of the ventral visual stream (Simmons et al., 2007), while words denoting actions (e.g., *kick*, *pick*) engage the dorsal motor network (Hauk et al., 2004). In recent years, much has been done to understand the automaticity, flexibility and reliability of the link between action/perception and word meaning (Martin and Chao, 2001; González et al., 2006; Simmons et al., 2007; Hoenig et al., 2008; van Dam et al., 2012). The current study extends this body of literature by addressing the question of how distributed lexical-semantic features are *integrated* during word comprehension.

Although ample evidence for the link between word meaning and perception/action systems exists, the bulk of research in this field has reduced lexical-semantic information to one dominant modality (e.g., vision for *red* and action for *kick*). The motivation for focusing on single modalities is clearly methodological: by focusing on words with a clear association to one modality, good hypotheses can be generated for testing empirically. However, words clearly refer to items that are experienced through multiple modalities in the real world (e.g., a football is associated with both a specific visual form and a specific action), and embodied accounts of language have done little to address how multimodal information interacts during the processing of word meaning. The one exception to this rule has been the attempt to understand how lexical-semantic processing can be focused flexibly on information from one modality versus another. For example, van Dam and colleagues (2012) demonstrated that words denoting objects that are strongly associated with both action and visual information (e.g., *tennis ball*) reliably activate both motor and visual pathways in the cortex. Interestingly, motor pathways also responded more strongly when participants were asked to indicate what to do with the object rather than what it looks like. Likewise, Hoenig and colleagues (Hoenig et al., 2008) have shown that even for objects with dominant modality-specific features (e.g., actions for artifacts), the pattern of activation in visual and motor networks is differentially modulated if a dominant (action) or non-dominant (visual) feature is primed. Notably, modality-specific networks show a stronger response to the target if the prime was not a dominant feature. Taken together, the studies by van Dam et al. (2012) and Hoenig et al. (2008) suggest that word meaning is partially stored in a network of areas that are recruited in a modality-specific and flexible way. However, it should also be pointed out that most of this evidence is of a correlational nature. As yet,

little is known about the causal role of modality-specific networks in lexical-semantic processing, and how they are related to more abstract semantic knowledge (Chatterjee, 2010; Hauk & Tschentscher, 2013).

While studies highlighting the flexible recruitment of different types of modality-specific information confirm that single words are associated with multiple types of perceptual experience, it is still unknown how information from multiple sources in the brain (e.g., visual and action features) is united to form a coherent concept that is both visual and motoric. Cross-modal integration has been studied extensively with respect to object perception (Schneider et al., 2008a, 2008b, 2011; Senkowski et al., 2008). However, its role in forming lexical-semantic representations has been largely neglected, even within the embodied framework. Several theoretical perspectives have argued for the existence of amodal integration ‘hubs’ or foci, at which information relevant for lexical-semantic processing is combined (Damasio, 1989; Patterson et al., 2007). Neuropsychological data has provided compelling evidence that the anterior temporal lobes (ATL) may be a good candidate for such a hub (Warrington, 1975; Hodges et al., 1992; Patterson et al., 2007). Thus, there is a general acceptance that information from distributed modality-specific networks is integrated in some way, somewhere in the brain. However, virtually no research has looked at what the neural mechanisms underlying semantic integration might be in these hub regions or more widely across the brain.

One way to investigate the mechanisms underlying integration across cortical areas is to study modulations in oscillatory power in EEG and MEG signals that have been related to network interactions at different cortical scales (von Stein and Sarnthein, 2000; Donner and Siegel, 2011). Specifically, low frequency modulations (< 20Hz) are often reported when tasks require the retrieval and integration of information from distant cortical sites, which is generally the case for memory and language (Bastiaansen, van der Linden, Ter Keurs, Dijkstra, & Hagoort, 2005; Bastiaansen, Oostenveld, Jensen, & Hagoort, 2008; Klimesch, 1999; Klimesch et al., 2010; Tallon-Baudry & Bertrand, 1999). In contrast, modulations in high frequency bands (>30Hz) are observed when tasks require local, modality-specific, network interactions such as saccade planning or visual object binding (Tallon-Baudry and Bertrand, 1999; Van Der Werf et al., 2008). According to this framework, the specific network dynamics underlying the integrating of lexical-semantic features across different modalities should be reflected in a modulation in low frequencies.

The aim of the current study was to investigate what mechanisms underlie the integration of semantic features across modalities. This question was addressed in two experiments using a dual property verification task. Participants were asked to indicate

whether a feature pair (e.g., *silver, loud*) is consistent with a target word (e.g., *WHISTLE*). Critically, the feature pair could either be from the same modality (e.g., both visual), or from different modalities (e.g., visual and auditory). In Experiment 1 we analyzed verification times for cross-modal and modality-specific feature contexts to investigate whether integrating multimodal semantic content, that is content, which is represented in distributed semantic networks, incurs a processing cost. Specifically, we hypothesize that integrating features represented within a single modality-specific network is faster than integrating features across modalities. In Experiment 2, we used EEG to measure changes in oscillatory neuronal activity during the target word when participants were asked to integrate features from the same or different modalities. Oscillatory neuronal activity could be a neural mechanism that contributes to semantic integration by linking modality-specific networks to multimodal convergence zones such as ATL. In line with this idea, we hypothesize that integrating semantic information from multiple modalities will be reflected in enhanced low frequency oscillatory activity in multimodal convergence zones, as well as substantial network interaction between these regions and a widespread cortical network.

Experiment 1

In Experiment 1 participants indicate whether two features (e.g., *silver, loud*) are consistent with a target word (e.g., *WHISTLE*). Specifically, a feature pair could either be associated with modality-specific or cross-modal semantic content. We hypothesize that integrating modality-specific feature pairs is faster than integrating cross-modal feature pairs, highlighting that word meaning is integrated more readily within modality-specific semantic networks than across.

Methods

Participants

Sixteen healthy individuals participated in Experiment 1 (13 female), all of which had normal or corrected to normal vision and no known auditory deficit. The age range was 18 to 24 ($M=19.88$). All participants were students at the University of York, and participated on a voluntary basis. As compensation for their participation, participants received either a financial reward or course credits. Participants gave written informed consent according to the Declaration of Helsinki. In addition they were given the opportunity of a more detailed debriefing after the study. The study was approved by the Ethics Committee of the Psychology Department at the University of York.

Stimulus material

120 target nouns (e.g., *WHISTLE*) were each paired with two adjective features from the same (e.g., *silver-tiny*), and two features from different modalities (e.g., *silver-loud*) (Figure 3.1A). Crucially, targets were presented only in one of the two feature contexts. That is, each participant saw 60 targets with a modality-specific (MS) feature pair and 60 different targets with a cross-modal (CM) feature pair. The conditions were counterbalanced and trials were presented in a pseudo-randomized order. In addition, 60 trials were included in which at least one feature was false. To familiarize participants with the experiment 10 additional practice trials were presented before the start of the experiment. Thus, each participant saw 190 target words and feature pairs.

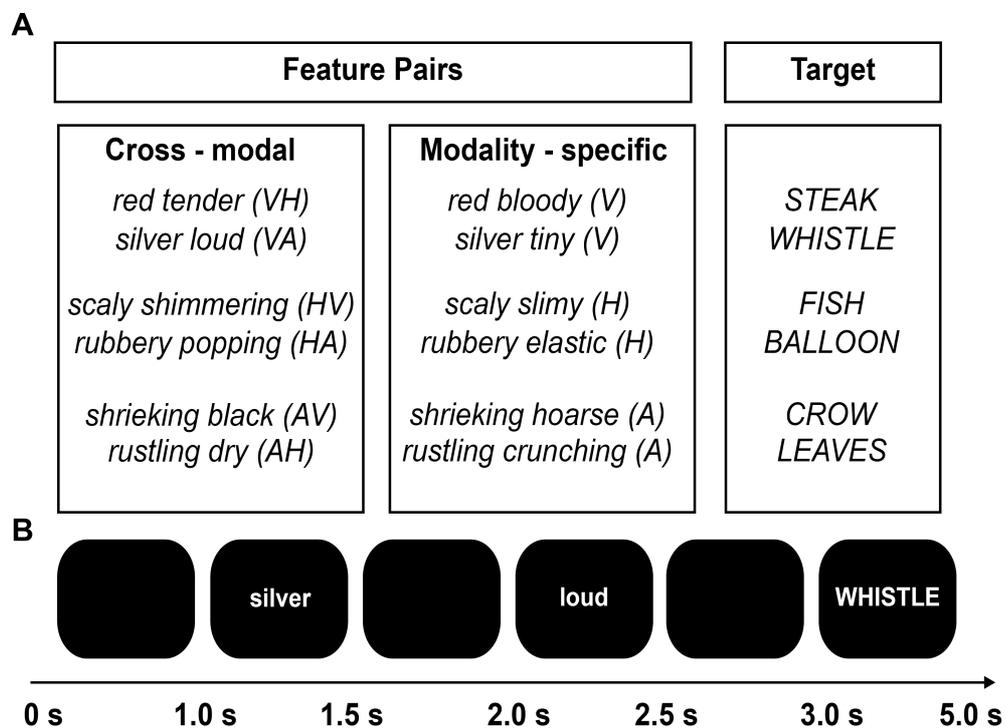


Figure 3.1. Experimental design of the dual property verification paradigm. A The top panel provides an overview of the design in which a target was either paired with a cross-modal (visual-haptic [VH; HV], visual-auditory [VA; AV], auditory-haptic [AH; HA]), or modality-specific feature pair (Visual [V], Auditory [A], Haptic [H]). The three modalities of interest were visual, haptic, and auditory. B The bottom panel depicts the time course of a single trial. All words are presented one after the other. Therefore, features can only be fully integrated when the target appears (e.g., *WHISTLE*).

Since the target (*WHISTLE*) and one feature (*silver*) were the same in both conditions, only variable features (*tiny - loud*) were matched for word frequency (log-scaled, British

National Corpus), and length. In order to control for differences in semantic association between feature pairs and targets, latent semantic analysis (LSA) scores were extracted for each feature pair and target combination. LSA is a measure of semantic similarity that quantifies how commonly two or more words occur in the same context in written texts (Landauer et al., 1998). For example highly associated words like *Camel* and *hump* yield a higher LSA score (LSA = .53) than less highly associated words such as *Camel* and *hairy* (LSA = .20). Lastly, each feature pair was rated on a five-point scale (N = 18) for how diagnostic and how related it is to its target word. None of these scores differed significantly between conditions (see Table 3.1).

Table 3.1. *Matching of the experimental items. Scores were averaged over all items in each condition. P-values were computed using independent-samples t-tests. The standard error of the mean is provided in brackets.*

Feature Pair	LSA	Relatedness	Diagnosticity	Frequency	Length
Cross-modal	0.21 (.01)	3.28 (.06)	2.65 (.07)	3.88 (.07)	6.48 (.18)
Modality-specific	0.22 (.01)	3.37 (.07)	2.66 (.08)	3.87 (.08)	6.24 (.17)
p-value	n.s. (p = .66)	n.s. (p = .32)	n.s. (p = .92)	n.s. (p = .93)	n.s. (p = .48)

Language is inherently polysemous, and most semantic features can be associated with multiple modalities, depending on the context. For example, a feature like high can be used to describe the size of a mountain (visual) or the pitch of a sound (auditory). This issue was addressed recently in two norming studies (Lynott and Connell, 2009; van Dantzig et al., 2010). Specifically, participants were asked to rate features in isolation or as feature-concept pairs on how likely the feature is experienced through one of five modalities (visual, haptic, auditory, olfactory, and gustatory). The features in the current study were based on averaged ratings from previous studies (Lynott and Connell, 2009; van Dantzig et al., 2010) and a small proportion (2.6%) of additional auditory features (e.g., *ticking*, *quacking*). Features were selected, which had been categorized as predominantly visual, haptic, or auditory (see Figure 3.2). All stimuli were presented using Neurobehavioral Systems Presentation software (www.neurobs.com) on a 22" TFT screen with a screen resolution of 1680x1050 and a refresh rate of 60 Hz.

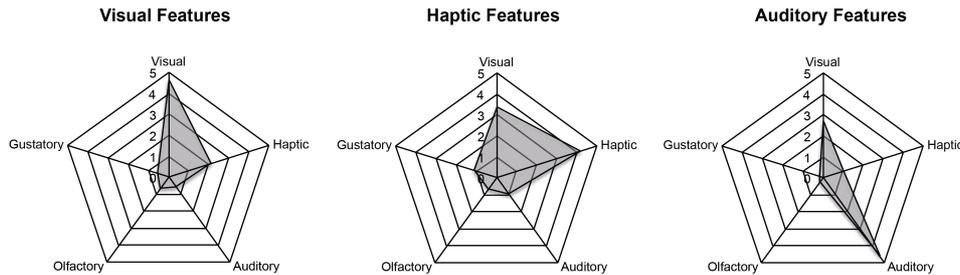


Figure 3.2. Mean of the modality ratings for visual, haptic, and auditory features. The three spider plots indicate the mean rating score (Lynott and Connell, 2009; van Dantzig et al., 2010) over all features in the each of the three modalities of interest (Visual, Haptic, and Auditory).

Procedure

Participants were seated in front of a computer screen at a distance of 40 cm. Words were presented in light grey on a black background with a font size of 40 pt. Each trial started with the disappearance of a fixation cross that was presented at a variable interval between 1500 and 2500ms. Individual features were presented subsequently, for 500ms, with a 500ms blank screen in between. The target was presented last (Figure 3.1B). Participants were instructed to indicate whether both features are consistent with the target. Responses were provided on a button box while the target was on the screen (2000ms). Response times and number of errors were measured for subsequent analyses. Each participant saw a target only once and in one of two conditions (CM or MS).

Results and Discussion

One participant was excluded from the analysis because performance rates on the task were at chance. Furthermore, outliers at three standard deviations from the mean were excluded from the analysis.

In order to test whether participants were able to perform the task, a one-sample t-test was conducted on the proportion of correctly identified feature-target pairs, against a test-value of 0.5. This test confirmed that participants' performance on the task was well above chance ($t(14) = 15.43, p < .001$) with a mean proportion of .73 correctly recognized features.

To test for a main effect of modality-specificity, the median reaction time was computed for each condition and participant, and averaged separately for MS (Visual, Auditory, Haptic) and CM (Visual-Auditory, Auditory-Haptic, and Visual-Haptic)

feature pairs, resulting in two values per participant (CM and MS). The distribution of these values across participants met the assumptions of a paired-sample t-test. The test statistic revealed that participants were overall slower to respond to CM ($M=981.6$, $SE=64.64$) versus MS ($M=909.36$, $SE=55.95$) feature pairs ($t(14)=3.65$, $p=.003$).

The effect of modality-specificity on verification time was further investigated for each of the three possible modality combinations using analysis of variance (ANOVA) with repeated measures (Figure 3.3). In each analysis, a CM condition (e.g., Visual-Auditory) was compared to two MS conditions (e.g., Visual and Auditory). The first ANOVA tested for an effect of condition on verification time across the visual (V), auditory (A), and visual-auditory (VA) conditions. The test revealed a significant main effect of condition (Wilks' Lambda = .33, $F(2,13)=13.24$, $p=.001$, partial $\eta^2=.67$). Planned comparisons using a Helmert contrast indicated that participants responded more slowly during CM (visual-auditory) than MS feature pairs (visual and auditory, respectively) ($F(1,14)=26.67$, $p<.001$, partial $\eta^2=.66$). The second ANOVA tested for a main effect of condition on verification time across the auditory (A), haptic (H), and auditory-haptic (AH) conditions. The results showed a significant main effect of condition (Wilks' Lambda = .43, $F(2,13)=8.61$, $p=.004$, partial $\eta^2=.57$). Planned comparisons using a Helmert contrast revealed that participants verified CM feature pairs (auditory-visual) more slowly than MS feature pairs (auditory and haptic respectively) ($F(1,14)=9.22$, $p=.009$, partial $\eta^2=.40$). The final ANOVA was conducted to test for a main effect of condition across the visual (V), haptic (H) and visual-haptic (VH) condition. There was no main effect in this analysis (Wilks Lambda = .72, $F(2,13)=2.64$, $p>.1$).

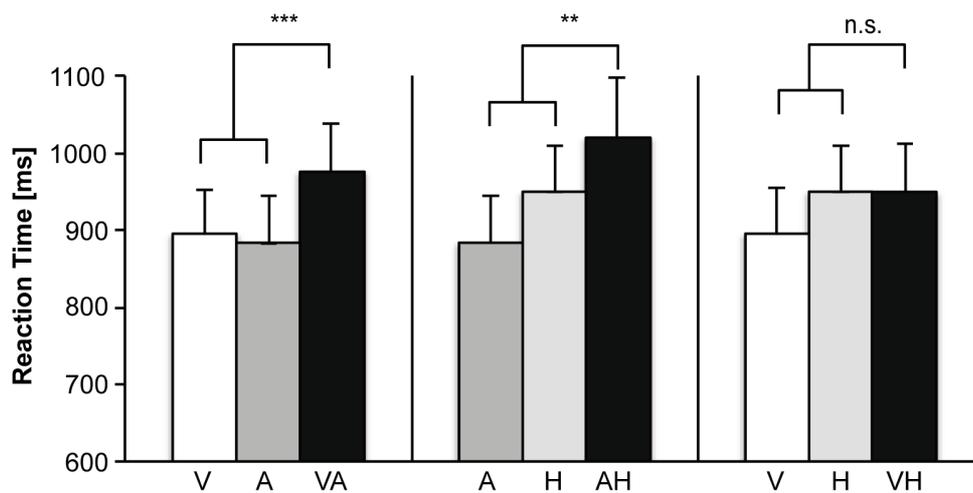


Figure 3.3. Cross-modal integration costs in verification times. Bar graphs depict the mean verification time in the MS (Visual, Auditory, and Haptic), and CM condition (Visual-Auditory, Auditory-Haptic, Visual-Haptic). Error bars denote standard error of the mean (** $p < .001$; ** $p < .01$).

The goal of Experiment 1 was to investigate whether integrating semantic features represented within a single modality is faster than integrating features across modalities. The current results suggest that this is indeed the case. Verification times for two semantic features with respect to a target (e.g., *WHISTLE*) were delayed when participants saw two features from different modalities (e.g., *silver, loud*). However, this effect seems to be restricted to visual-auditory, and auditory-haptic feature combinations. A possible explanation for this finding is that visual lexical-semantic features can be difficult to distinguish from haptic features. This was also evident in the rating study in which features were often rated similarly as being experienced by seeing, and touching (Figure 3.2) (Lynott and Connell, 2009; van Dantzig et al., 2010).

Experiment 2

Experiment 2 uses EEG to investigate oscillatory dynamics during semantic integration within, and across different modalities. We hypothesize that integrating cross-modal semantic content will be reflected in enhanced low frequency oscillatory activity in multimodal semantic hubs, such as ATL, as well as substantial network interaction between these regions and a widespread cortical network.

Methods

Participants

For Experiment 2, 22 healthy participants (8 female) were tested, all of which had normal or corrected to normal vision and no known auditory deficit. The age range was between 19 and 34 ($M=21.26$). Four participants were excluded from the analysis due to excessive movement and blinking (3), and a technical error (1). None of the participants had participated in Experiment 1. Participants gave written informed consent according to the Declaration of Helsinki. In addition they were given the opportunity of a more detailed debriefing after the study. The study was approved by the Ethics Committee of the Psychology Department at the University of York.

Stimulus material

The stimulus materials in Experiment 2 were exactly the same as in Experiment 1.

Procedure

In Experiment 2, participants were wearing an electrode cap that was connected via an amplifier to the recording computer while performing the verification task. The setting was the same as in Experiment 1. However, in order to prevent contamination of the EEG signal from movement and response planning (Neuper et al., 2006), the task was changed such that participants only responded in case they encountered a false feature.

Data recording and pre-processing

EEG was acquired from 64 Ag-AgCl electrodes that were positioned on an electrode cap according to a 10-20 system. All electrodes were re-referenced offline to the algebraic average of the two mastoids. Horizontal and Vertical eye movements were recorded with a set of bipolar Ag-AgCl electrodes. The signal was amplified using an ANT amplifier with a band-pass filter between 0.5 and 100 Hz. Impedances of the cortical electrodes were kept below 10 k Ω . The signal was recorded with a sampling frequency of 500 Hz.

Offline analyses were conducted using Matlab 7.14 (Mathworks, Natick, MA) and Fieldtrip, a Matlab toolbox for analyzing EEG/MEG data (Oostenveld et al., 2011). Trials were only considered if the participant correctly withheld the response on a target. Artifact rejection was performed in three consecutive steps. First, muscle artifacts were removed using semi-automatic artifact rejection. Subsequently, extended infomax independent component analysis (ICA), with a weight change stop criterion of $< 10^{-7}$, was performed to identify, and reject ocular components. Finally, each trial was visually

inspected for any remaining artifacts. The average number of correct trials that survived the rejection protocol did not differ significantly between condition (MS: $M=48$, $SE=1.26$; CM: $M=47$, $SE=1.24$; $t(17)=-1.29$, $p=.21$).

Spectral analysis

In order to estimate spectral power changes over time, time-frequency representations (TFR) were computed for each trial, using a 500 ms fixed sliding time window with time steps of 50 ms, resulting in a frequency resolution of ~ 2 Hz. A Hanning taper was applied to each of these segments to reduce spectral leakage. TFR's were calculated for frequencies between 2 and 20 Hz in steps of 2 Hz. These transformations were performed at the individual trial level and reflect both evoked and induced components of the signal. Subsequently, trials were averaged for each condition and subject, and percentage signal change was computed using a common baseline over both conditions. The time window for the baseline was between 750 and 250 ms before the onset of the trial. The baseline normalization procedure is equivalent to the event-related de-synchronization technique (Pfurtscheller & Lopes da Silva, 1999), except that positive values denote synchronization, and negative values de-synchronization³. Total power was averaged over 6 regions of interest (Figure 3.5).

Statistical analysis

Inferential statistics on the time-frequency windows following the presentation of the target word were computed using a cluster-based permutation approach (Maris and Oostenveld, 2007). Cluster-based permutation effectively reduces the number of comparisons by clustering neighboring samples above a given threshold along the dimensions: time, frequency, and space. In the current study, paired-sample t-tests were computed over subjects for each ROI-time-frequency point (0-1000 ms, 2-20 Hz, 6 ROI). Subsequently, t-values were thresholded at $\alpha=.05$. Neighboring t-values above the threshold criterion were included into the same cluster, and ranked according to the size of the cluster. Finally, cluster-level statistics were computed by comparing the sum of all

$$3 \left(\frac{\text{activePeriod} - \text{passivePeriod}}{\text{passivePeriod}} \right) * 100$$

t-values within a given cluster against a permutation null-distribution. The null-distribution was constructed by randomly permuting the conditions (iterations=1000), and calculating the maximum cluster-level statistic for each iteration.

A similar procedure was used for the seed-based whole-brain connectivity analysis. The difference between each condition (CM and MS) and the baseline was computed for an early (0-500 ms) and late (500-1000 ms) time window. The value at each location in source space was thresholded using a permutation distribution ($\alpha=0.05$, 1000 iterations), and combined with values from spatially adjacent locations. We used a maximum statistic to control for multiple comparisons at the cluster-level, which was equivalent to the sensor space analysis.

Source reconstruction

The forward model was computed using the Boundary Element Method (BEM) on a regular three dimensional grid (10x10x10mm spacing), containing 3000 vertices. The model is based on the segmentation of the colin27 template and is described in detail in Oostenveld, Stegeman, Praamstra, & Van Oosterom (2003). As the same model was used for all individuals, no additional normalization of individual grids was required.

Sources of oscillatory activity at the whole-brain level were estimated using a frequency domain beamforming method called Dynamic Imaging of Coherent sources (DICS: Gross et al., 2001; Liljeström, Kujala, Jensen, & Salmelin, 2005). DICS uses an adaptive spatial filter, which is a linear transformation, that when applied to the data passes activity at each source location with unit gain, while suppressing all other sources. The transformation matrix is computed as a constrained optimization problem on the basis of the cross-spectral density matrix (CSD) and the solution to the forward model (lead field). The CSD matrix was computed between all 30 EEG electrodes in the time and frequency window of the effect of interest (500-1000ms, ~4-8Hz) using a Hanning taper. In addition, a regularization of 5% was introduced to avoid non-invertibility of the CSD. The power in the dominant source direction was computed using single value decomposition.

Paired-sample t-tests were computed for the difference between conditions at each location in the brain. Subsequently, t-values were transformed into z-values and masked at $\alpha = 0.05$.

Connectivity analysis

The analysis of cortico-cortical connectivity in source space was conducted for an early (0-500 ms) and a late time window (500-1000 ms) at the frequency that showed the strongest power difference in sensor space (~6 Hz). The same number of trials were

randomly selected for the CM and MS condition as well as the baseline period. A cross-spectral density (CSD) matrix was computed from the tapered Fourier spectra of each trial and used to estimate filter coefficients for the adaptive spatial filter. Subsequently, the Fourier spectra were projected through these filter coefficients along the strongest dipole orientation.

Functional connectivity between each location in the brain and all others was estimated using the imaginary part of coherency (ImCoh). ImCoh is only sensitive to signals at a non-zero time-lag, and therefore insensitive to connectivity artifacts resulting from volume conduction (Nolte et al., 2004). We computed ImCoh based on the Fourier spectra at each location in the grid. Subsequently, a stabilizing z-transform was applied using the inverse hyperbolic tangent (\tanh^{-1}). Since the main interest was in the functional connectivity between nodes rather than the direction of the effect, the absolute was computed for each of the resulting z-values.

For subsequent graph analysis, a binary adjacency matrix was computed for each participant by thresholding with the maximum value at which none of the nodes in any of the conditions was disconnected to the rest of the network. Finally, the log10 transformed difference between the number of connections (degrees) in the seed region versus baseline was computed for each condition, and subjected to statistical testing.

Results and Discussion

The time-frequency analysis of total power revealed a sustained increase in the theta band (4-6 Hz) and a decrease in the alpha, and low beta band (8-20Hz) while the target word (e.g., *WHISTLE*) was on the screen (Figure 3.4A). In order to test for differences between conditions (CM>MS), a cluster-based permutation approach was used (Maris and Oostenveld, 2007). In the first step of the analysis, the clustering algorithm revealed one significant cluster (4-6 Hz, peak at 750-850 ms) at left and central electrodes (LA, LP, MA, MP) (Figure 3.4B; Figure 3.5). In order to control for multiple comparisons, a maximum permutation statistic was used in which the summed cluster t-value was compared against a permutation distribution with 1000 iterations. The maximum statistic revealed a significant difference between conditions at the cluster level ($p=.002$, two-tailed), suggesting enhanced theta power in the cross-modal condition. Source reconstruction of this effect revealed a major peak in left ATL as well as left middle occipital gyrus (MOG) (Figure 3.6A).

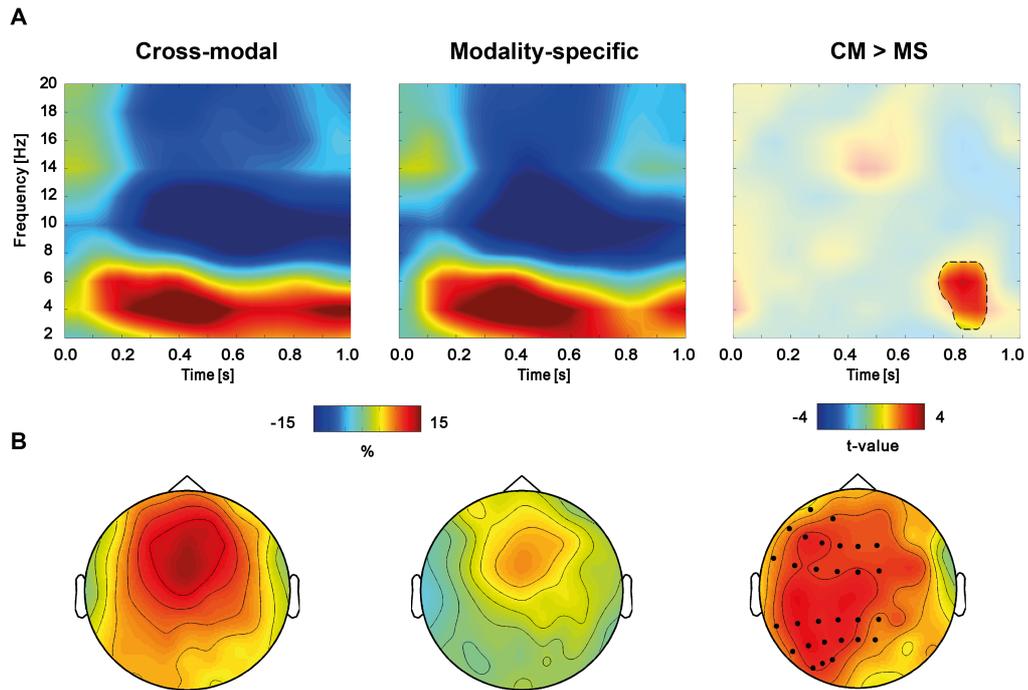


Figure 3.4. Modulation in low frequency cortical oscillations for the target word in a cross-modal or modality-specific context. A The top panel shows time-frequency representations, averaged over all significant clusters. The first two panels show the grand average percent signal change with respect to the baseline. The third panel depicts the masked statistical difference between the two conditions in t-values. The contour plot reveals one significant cluster in the theta range (4-6 Hz). B The first two bottom panels depict the topography of the effect in each condition (4-6 Hz, peak at 750-850 ms) relative to baseline. The third panel signifies the statistical difference between conditions in t-values. Electrodes within significant clusters are marked with dots ($p=.002$, cluster-corrected)

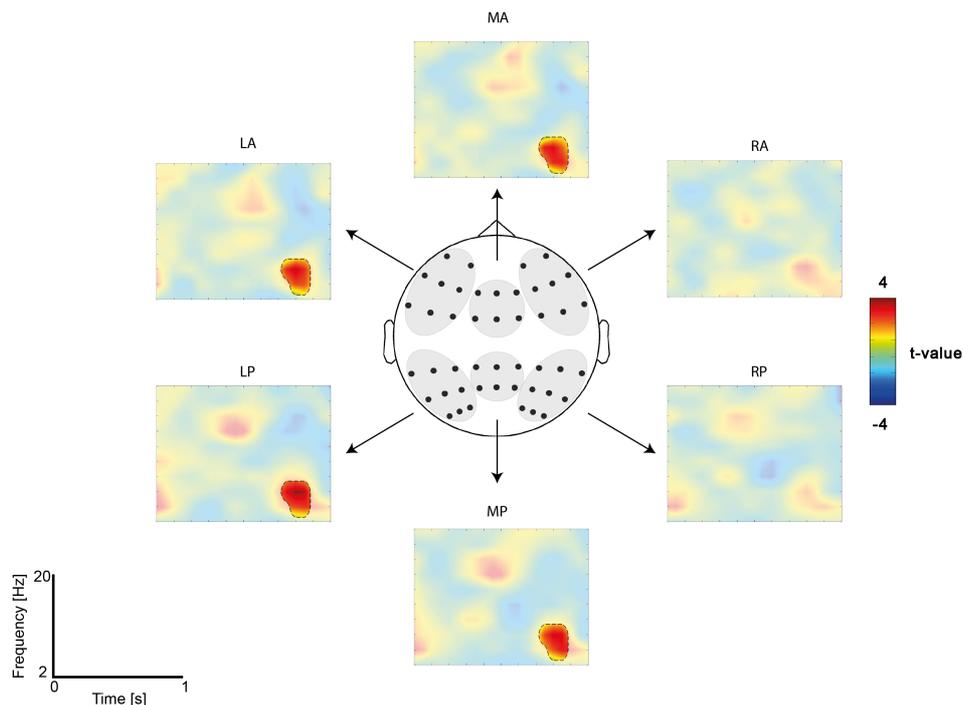


Figure 3.5. Time-frequency plots for each of the 6 ROI. The ROI were middle anterior (MA), left anterior (LA), right anterior (RA), middle posterior (MP), left posterior (LP), and right posterior (RP) electrodes. Time-frequency representations depict the statistical difference in t-values for the target word in the CM versus MS feature context. The contours indicate the peak of the cluster-corrected statistical difference ($p=.002$).

The grid point in the left ATL (mni coordinate: -49 22 -30), which was most sensitive to the power difference between conditions, was taken as the seed for subsequent connectivity analyses. One sample t-tests were used to test for an increase in the log-transformed number of connections (degrees) relative to baseline in an early (0-500 ms) and late (500-1000 ms) time window. In the early time window, both conditions showed a significant increase in the number of connections (CM: $t(17) = .389$, $p<.001$; MS: $t(17) = .355$, $p=.001$, one-sided). However, in the late time window, an effect was found only in the CM condition (CM: $t(17) = 2.13$, $p=.024$; MS: $t(17)=.56$, $p=.291$, one-sided). Further, paired-sample t-tests were used to test for a difference between conditions directly. A difference between conditions was observed only in the late ($t(17)=2.36$, $p=.031$, two-sided), but not the early time window ($t(17)=.012$, $p=.991$, two-sided). Taken together, this suggests that during the first 500 ms after target onset, the ATL is communicating with a wide cortical network in both conditions (CM and MS). However, during the second 500 ms, this effect persists only in the CM condition (Figure 3.6B).

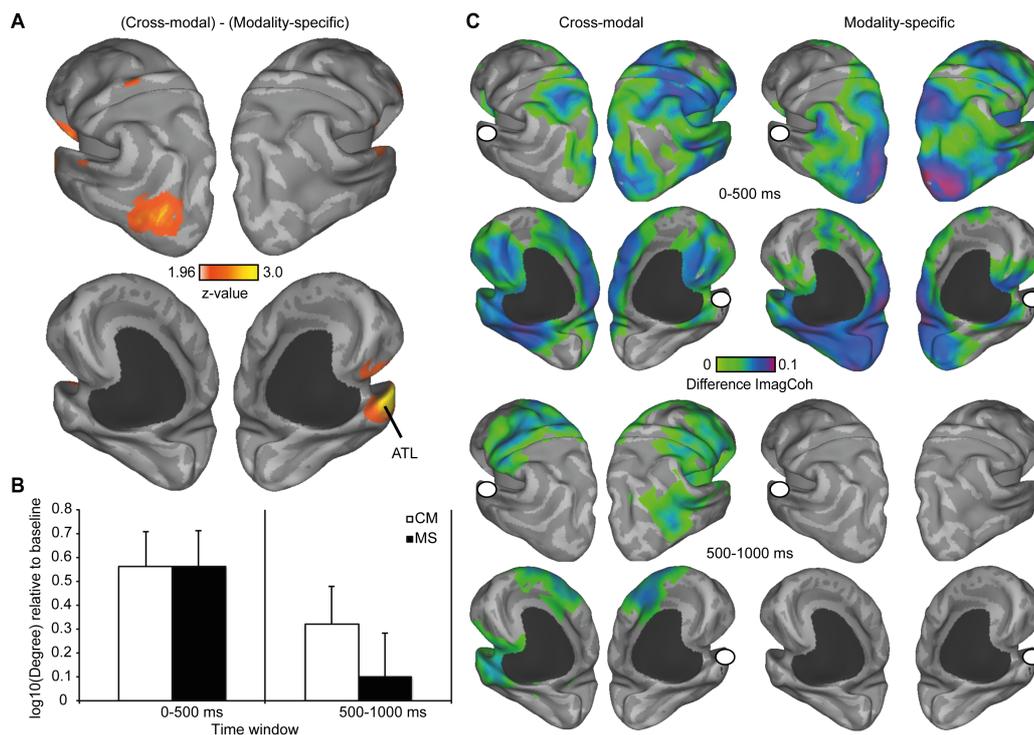


Figure 3.6. Source reconstruction and connectivity analysis. **A** Source reconstruction of the effect in the theta band, depicted as thresholded z-values, reveals peaks in left ATL and MOG. **B** Bar graphs show a significant increase in the number of connections between ATL and the rest of the brain in the early time window (0-500 ms). In the late time window (500-1000 ms), only the CM condition shows a significant increase in the number of connections relative to baseline. Error bars depict SEM **C** Results of the whole-brain connectivity analysis, seeded in the ATL (white dot). Connectivity maps show the difference in absolute, z-transformed, imaginary coherence between each condition and the baseline. In the early time window both conditions show a strong increase in connectivity between the ATL and a widespread cortical network. In the second time window, only the cross-modal condition shows continuing network activity above baseline.

To illustrate which specific regions show enhanced functional connectivity with the ATL, we used a whole-brain cluster-based permutation procedure on the z-transformed ImCoh values, comparing each condition to the baseline. This approach was similar to the procedure we used in sensor-space. As depicted in the top panel of figure 3.6C, a large cluster of nodes was connected to the ATL in the early time window for both conditions (CM: $p=.004$; MS, $p=.008$, one-sided). However, in the late time window a significant difference relative to baseline was only observed in the CM condition ($p=.032$, one-sided). Connections during the second time window were found to regions

that are involved in auditory (right Heschl's gyrus), somatosensory (bilateral post-central gyrus), and visual object processing (right posterior MTG), as well as medial and lateral frontal lobes.

The aim of Experiment 2 was to investigate whether integrating semantic features over a wider cortical network is reflected in enhanced oscillatory activity at low frequencies. Time-frequency analysis revealed an increase in theta power (4-6 Hz) for both conditions, which was more sustained during cross-modal integration. This effect localized most strongly to the left ATL, which is thought to be a major hub for integrating multimodal semantic content (Patterson et al., 2007). Subsequent seed-based whole-brain connectivity analysis confirmed that the number of connections between the ATL and the rest of the network increases in both CM and MS conditions during the first 500 ms. However, these network interactions extend into the second 500 ms only in the CM condition. Specifically, the ATL communicates with modality-specific auditory, somatosensory and high-level visual areas as well as regions in the frontal lobe. Taken together, these findings suggest that theta oscillations are associated with network dynamics in a widespread cortical network. Previous research has associated theta oscillations with lexical-semantic processing (Bastiaansen et al., 2005; Bastiaansen et al., 2008). However, the current study is the first to show that theta power is sensitive to the spatial distribution of semantic features in the cortex. The implications of these findings for semantic processing are discussed in the next section.

General discussion

Embodied theories of language have argued that word meaning is partially stored in modality-specific cortical networks, converging in multisensory association areas in anterior temporal, and inferior parietal lobes (Barsalou, 2008; Binder & Desai, 2011; Damasio, 1989; Pulvermüller & Fadiga, 2010; Vigliocco et al., 2009). The aim of the current study was to investigate the mechanisms underlying integration of semantic features during language processing. Two experiments are reported in which participants were asked to verify whether two features from the same (e.g., *silver - tiny*) or different modality (e.g., *silver - loud*) are consistent with a given target word (e.g., *WHISTLE*). The results from Experiment 1 show that integrating features from the same modality is faster than integrating features from different modalities. These findings suggest that word meaning is integrated more readily within a single modality-specific network than across networks. Integrating information across networks in particular should engage multimodal convergence zones. Experiment 2 shows that integrating features from different modalities induces a sustained theta power increase in left ATL, a putative hub

for semantic convergence (Patterson et al., 2007). Low frequency theta oscillations could reflect a neural mechanism by which multimodal word meaning is combined locally in temporal association cortices. However, assuming that word meaning is partially stored in distributed cortical networks, multimodal integration necessarily requires long-range communication between left ATL and the rest of the cortex. The seed-based connectivity analysis in the theta range revealed that this is indeed the case; left ATL communicates with a widespread cortical network that includes, but is not limited to, modality-specific regions. In other words, local theta power in left ATL reflects long-range communication between temporal areas and the rest of the cortex, which, according to embodied theories of semantics, is necessary for the integration of word meaning from multiple modality-specific semantic networks.

Integrating multimodal semantic information comes at a cost

Experiment 1 shows that participants are faster to verify features of a target word (e.g., *WHISTLE*) from the same (e.g., *silver-tiny*) versus two different modalities (e.g., *silver-loud*), suggesting that word meaning converges more readily within a modality-specific semantic network than across networks. This is in line with behavioral studies that have examined switching costs during word comprehension (Pecher et al., 2003) as well as dual property verification tasks (Barsalou et al., 2005) (but also see McNorgan, Reid, & McRae, 2011). It is also broadly in accordance with a cognitive model proposing graded semantic convergence from modality-specific to multimodal (Plaut, 2002).

Theta oscillations in left ATL during multimodal semantic feature integration

The principle by which information from distributed neural populations is combined is a much-debated topic in neuroscience. It has been argued that transient networks emerge from synchronized firing of large neuronal populations, which is recorded as oscillatory activity at the scalp (Milner, 1967; von der Malsburg and Schneider, 1986; Singer and Gray, 1995). In humans, changes in oscillatory neuronal activity in the theta range have been observed during different stages of memory processing, as well as lexical-semantic retrieval (Bastiaansen et al., 2005; Bastiaansen et al., 2008; Jensen & Tesche, 2002; Klimesch, 1999; Klimesch et al., 2010; Raghavachari et al., 2006; Summerfield & Mangels, 2005; Wu et al., 2007). The current study extends previous findings to show that theta oscillations are particularly sensitive to the *integration* of semantic features of an object, which are thought to be partially represented in distributed, modality-specific, networks (Barsalou, 2008; Vigliocco et al., 2009; Pulvermüller and Fadiga, 2010; Binder and Desai, 2011).

It has been argued that modality-specific semantic networks converge in multimodal association cortices (Damasio, 1989; Patterson et al., 2007). For example, there is compelling evidence from patients with semantic dementia suggesting that ATL is involved in semantic processing at a general level (Warrington, 1975; Hodges et al., 1992; Patterson et al., 2007), yet little is known about the neural dynamics within this region. Experiment 2 reports a modulation in local theta power within left ATL when participants integrate features from multiple modality-specific semantic networks. The connectivity analysis of the data from Experiment 2 further revealed that theta oscillations also participate in long-range network interactions linking left ATL with a widespread cortical network. These findings are an important step in bridging the gap between anatomy and cognition; the theta rhythm could be a neural signature reflecting transient network interactions within left ATL, as well as between this region and distributed modality-specific networks. Such functional networks are necessary for linking semantic content in space and time.

Lastly, we find that the effect peaks at a very late point in time (~750 ms), most likely reflecting the tail of a sustained oscillatory response that is triggered much earlier in time. Importantly, we do not argue that this is the moment when semantic integration takes place. Rather, oscillatory dynamics in the theta range could be involved in creating the conditions necessary for semantic integration by linking multiple functional networks over a period of time. The fact that cross-modal integration incurs a higher processing demand is reflected in a longer integration window. This is also in line with the finding that theta is the only known oscillatory frequency which shows a linear increase during sentence processing (Bastiaansen et al., 2002). Again, we would like to emphasize that the primary goal of the current study was to investigate the oscillatory dynamics, rather than the timing of semantic integration, which has been addressed extensively in previous work using the event-related potential technique (Kutas and Federmeier, 2000).

Relation to multisensory integration and cross-modal matching

Multisensory integration is an essential component of everyday life. For example, both visual and proprioceptive information are required when performing goal-directed actions (Sober and Sabes, 2003), speech comprehension greatly benefits from visual information about lip movements (Rosenblum, 2008), and hearing the sound of an animal facilitates its visual detection (Schneider et al., 2008b). Although these examples bear a superficial resemblance to the processes we investigated in the current study, it should be noted that there are fundamental differences in integrating cross-modal

sensory, and lexical-semantic content respectively. These differences are with respect to a) the time scale and b) the directionality of information flow.

Previous studies have investigated oscillatory changes during multisensory integration using cross-modal matching. For example, Schneider and colleagues (Schneider et al., 2008a) showed that matching the visual image of an object (e.g., picture of a sheep) to its sound (e.g. sound of a sheep) induces an early increase in the gamma band (40-50 Hz) between 120-180 ms. Similar findings have been reported for haptic to auditory matching (Schneider et al., 2011). In contrast, effects of semantic integration in language are usually observed around 400 ms (Kutas and Federmeier, 2000) and at frequencies below 30 Hz (Bastiaansen et al., 2005; Bastiaansen et al., 2008; Hald et al., 2006; Wang, Jensen, et al., 2012) (however, see Hagoort, Hald, Bastiaansen, & Petersson, 2004). This is not surprising given the fact that lexical retrieval involves multiple processing stages (e.g., visual processing of letters). In this respect, the current findings should primarily be interpreted as reflecting language but not sensory processing.

Another difference between sensory and semantic integration is the directionality of information flow. While sensory processing in a given modality is largely automatic and dependent on external stimulation (bottom-up), retrieving modality-specific word meaning requires prior experience with the referent of a word and is highly context-dependent (top-down). For example, previous imaging work has shown that action words do not activate the action system to the same extent if they are presented as idiomatic expressions (e.g., *he kicked the bucket*) (Raposo et al., 2009) (but, see Boulenger, Hauk, & Pulvermüller, 2009). Furthermore, it has been shown that neutral sentences (e.g., *it is hot in here*) activate parts of the action system if presented in a context in which they are interpreted as indirect requests (e.g., a room with a closed window) (van Ackeren et al., 2012). In the current study, participants were primed to think about a particular instance of an object (e.g., *a silver and loud whistle*). In other words, the relevant information was not directly encoded in the stimulus (a visual word), but needed to be retrieved from memory.

In sum, imaging studies have shown that lexical-semantic content activates modality-specific cortical networks similar to sensory stimulation (Hauk et al., 2004; Kiefer et al., 2008; Simmons et al., 2007). But despite their spatial similarity, lexical-semantic and sensory processes operate at very different time-scales and through different computations (bottom-up versus top-down). While much is known about the mechanisms underlying multisensory integration, the current study is among the first to address how cross-modal semantic information is integrated through language.

Conclusions

Previous research suggests that word meaning is partially stored in modality-specific cortical networks. However, little is known about the mechanisms by which distributed semantic information is combined into a coherent conceptual representation. The current study addresses exactly this question: What are the mechanisms underlying cross-modal semantic integration? Participants were asked to indicate whether two features from the same (e.g., *silver - tiny*) or different modalities (e.g., *silver - loud*) are consistent with a target word (e.g., *WHISTLE*). Experiment 1 revealed that integrating semantic features represented within a single modality is faster than integrating features across modalities. In Experiment 2, EEG recordings revealed sustained oscillatory activity in the theta range, when participants were asked to integrate features from different modalities. The effect was localized to left ATL, a putative semantic hub that is thought to be involved in linking multimodal semantic content (Patterson et al., 2007). While the importance of this region for semantic processing and integration is largely uncontested, little is known about its mechanics. The current findings are an important step towards bridging this gap between anatomy and function; oscillatory dynamics in the theta range could be a neural mechanism that is involved in establishing transient functional connections between distributed modality-specific, and multimodal semantic networks. Further evidence for this claim is the finding that theta oscillations in Experiment 2 also participate in long-range interactions linking left ATL to a widespread cortical network.

Chapter 4:

Oscillatory dynamics in the theta band reflect cross-modal feature integration not switching

Abstract

Research from the previous decade suggests that word meaning is partially stored in modality-specific cortical networks that are distributed across the cortical surface. Yet, little is known about how this information is combined into a coherent word representation. It has been suggested recently, that low frequency cortical oscillations in the theta band could reflect a mechanism involved in linking multiple memory networks (van Ackeren and Rueschemeyer, 2014). The aim of the current study is to test whether theta oscillations are indeed sensitive to the integration of multimodal semantic knowledge as argued by van Ackeren and Rueschemeyer or rather switching between different modalities (Pecher et al., 2002). A dual property verification task is used in which two features are combined with a target word. Word content of the feature pairs was either from the same or different modalities (modality-specificity), while the target words could be either congruent or incongruent with the features (congruency). EEG was recorded while participants performed the task to test for modulations in low frequency power (4-6 Hz). Statistical analysis of spectral power in the theta range revealed an interaction effect between the factors modality-specificity and congruency. Specifically, theta power was enhanced for cross-modal versus modality-specific feature target combinations. However this effect was only observed for congruent feature target combinations. Taken together, these results replicate the earlier finding of sustained theta oscillations during cross-modal feature verification. Furthermore, the present study corroborates the claim that low frequency theta oscillations are sensitive to cross-modal semantic integration rather than modality switching.

Introduction

Embodied theories of language argue that word meaning is partially stored in modality-specific cortical networks (Barsalou, 2008; Binder et al., 2009; Pulvermüller and Fadiga, 2010). For example, multiple studies have demonstrated that words denoting actions engage a cortical motor network (Hauk et al., 2004; Postle et al., 2008), and words denoting visual features (color, shape) engage visual networks (Simmons et al., 2007; Zwaan et al., 2002). While there is much evidence for the existence of these modality-specific networks, little is known about how information from multiple networks is integrated into a coherent conceptual representation during language understanding.

Over the last decade, most studies have tested predictions of embodied language theories using words from a single dominant modality (e.g., action words, colors, sounds) (Hauk et al., 2004; Kiefer et al., 2008; Simmons et al., 2007). However, more recent research stresses that words are associated with information from multiple modalities that is accessed in a dynamic, and context-dependent way. For example, van Dam and colleagues (2012) demonstrated that words with strong bimodal (visual and auditory) associations (e.g., *tennis-ball*) engage both action (inferior parietal lobule) and visual (inferior temporal lobe) association networks. Furthermore, the authors showed that the pattern of activation in these networks is modulated differently if participants are asked to think about visual- versus action-related features of the object. Converging evidence for the flexibility of these modality-specific networks has also been reported in a different study for words with one dominant modality (visual or motoric) (Hoenig et al., 2008). Taken together, these results illustrate that semantic information is at least partially represented in distributed functional networks that dynamically adapt to the demands of the task.

The distributed organization of semantic knowledge has direct implications for word processing. Specifically, as information transfer over long distances is costly (e.g. due to conduction delays), it should be easier to process information that is represented in spatially adjacent versus more distributed networks. Indeed, property verification studies have shown that participants are faster to verify semantic features of a target noun following features from the same versus different modality. Interestingly, this effect was also found for novel/impossible feature concept relations such as jingling onion (Connell and Lynott, 2011). Similarly, dual property verification studies have shown that participants are faster to evaluate two features of a target concept (e.g., whistle) from the same (tiny, silver) versus different modalities (tiny, loud) (Barsalou et al., 2005; van Ackeren and Rueschemeyer, 2014). These results support the idea that

processing spatially distributed information is more costly than processing modality-specific information.

The physiological mechanisms of how information is combined within and across modalities are still poorly understood. One possibility is that neuronal populations communicate by firing in synchrony (Singer and Gray, 1995; Womelsdorf et al., 2007), which is reflected in oscillatory activity during MEG/EEG recordings. For example, van Ackeren and colleagues (van Ackeren and Rueschemeyer, 2014) asked participants to indicate whether two features from the same or different modality are consistent with a target word, while recording oscillatory neuronal activity in EEG. The study revealed a more sustained increase in low frequency (theta) power in left anterior temporal lobe (ATL), a putative hub supporting multimodal semantic integration (Patterson et al., 2007). The same region showed enhanced communication with a widely distributed cortical network in the theta band when participants verified features from different modalities. The authors argue that slow cortical theta oscillations could be involved in linking modality-specific and supramodal semantic networks, which is critical for semantic feature integration. However, these results might also be attributed to a modality-switching effect in which information from two features is never fully integrated with the concept.

The aim of the current study is to test whether sustained low frequency theta power during cross-modal feature integration reflects feature integration, or modality switching. To this end, participants perform a dual feature verification task where features are presented from the same (auditory or visual) or different modalities (audio-visual). In addition, all feature combinations are paired with congruent, or incongruent target words. If the sustained theta response is sensitive to semantic integration, there should be a modality-specificity effect in the congruent, but not the incongruent conditions. In contrast, switching effects have been observed even for novel or incongruent feature-target (jingling onion) combinations. Thus a sustained theta response would be expected during cross-modal trials in both the congruent and incongruent condition if the effect reflects modality switching.

Methods

Participants

Participants were 32 healthy individuals (9 male) between 18 and 29 years ($M=22.11$), who were native speakers of English and had no known neurological impairment. All participants were students at the University of York, and participated on a voluntary basis. Participants received either a financial reward or course credits for taking part in

the experiment. All participants gave written informed consent according to the Declaration of Helsinki and were debriefed after the study. The study was approved by the Ethics Committee of the Department of Psychology at the University of York.

Stimulus material

304 target nouns (e.g., *APPLE*) were paired with two feature words from the same (e.g., *green - round*), and two features words from a different modality (e.g., *green - crunchy*). 152 nouns were paired with congruent feature pairs (modality-specific and cross-modal) and 152 nouns were paired with incongruent feature pairs. Thus, the same feature pair was used once in a congruent condition and once in an incongruent condition. In addition, each participant saw each noun only once, that is, in a cross modal or a modality-specific feature context.

Modality-specific (auditory or visual) features of the 152 nouns in the congruent condition were selected based on a previous experiment and a rating (N=46) in which participants were given a list of nouns and asked to provide two features that describe what the object sounds like (auditory) or looks like (visual). Subsequently, features (Table 4.1) and target nouns (Table 4.2) were matched across conditions for word frequency (Brysbaert et al., 2012), and length. In addition, a separate rating (N=20) was conducted to ensure that feature combinations in the congruent modality-specific and cross-modal condition were equally diagnostic with respect to the target noun. Specifically, participants were given a list of feature pairs and asked to predict the target word. The number of participants that correctly predicted the target word was matched across conditions. Thus, the target word was not more predictable in the cross modal versus modality-specific condition (Table 4.1).

Table 4.1. *Matching of features for psycholinguistic variables and diagnosticity.*

Condition	Word length	Word frequency (log10)	Diagnosticity
Cross-modal			
M	5.28	2.64	1.11
SE	0.13	0.07	0.12
Modality-specific			
M	5.30	2.62	1.07
SE	0.10	0.07	0.12
<i>p</i>	0.87	0.74	0.82

Table 4.2. *Matching of nouns in the congruent and incongruent condition.*

Condition	Word length	Word frequency (log10)
Congruent	M	6.04
	SE	0.19
Incongruent	M	6.07
	SE	0.17
	<i>p</i>	0.9

Procedure

Participants performed the task on a computer screen at a distance of 40cm. Words were presented in white on a black background and a font size of 40 pt. Before the start of the main experiment, participants were familiarized with the task during a practice run (10 trials). Each trial began with a fixation cross (1000ms), followed by the first and second feature word (800ms, each) with a short fixation period between features (200ms). The target word was presented for 2000ms. To avoid response related contamination of the EEG signal participants were instructed to respond only after the target word was replaced by a question (Figure 4.1). Responses were provided on a button box.

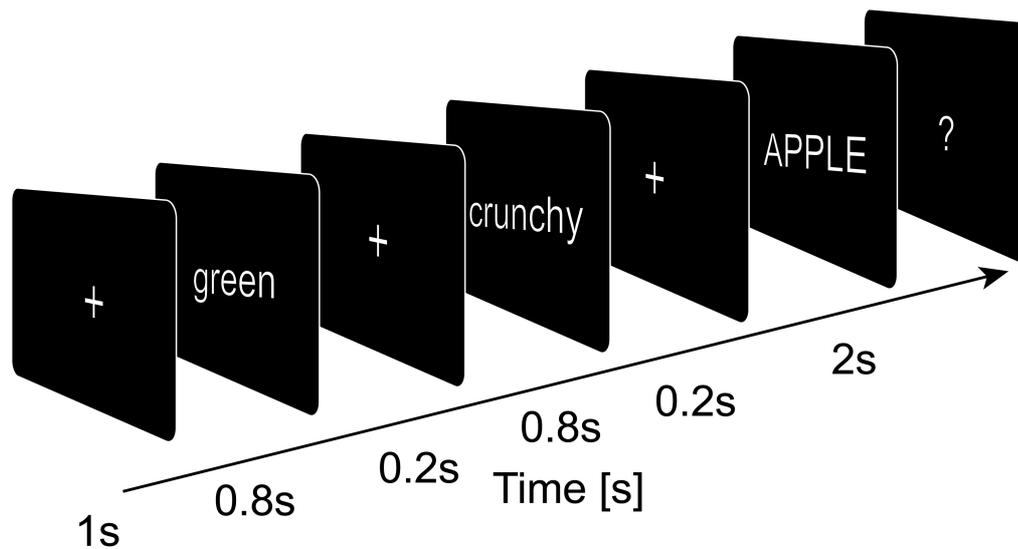


Figure 4.1. Timeline of the dual feature verification task.

Data recording and pre-processing

EEG was acquired from 64 Ag-AgCl electrodes that were positioned on an electrode cap according to a 10-20 system. All electrodes were re-referenced offline to the algebraic average of the two mastoids. Horizontal and Vertical eye movements were recorded with a set of bipolar Ag-AgCl electrodes. The signal was amplified using an ANT amplifier with a band-pass filter between 0.5 and 40 Hz. Impedances of the cortical electrodes were kept below 10 k Ω . The signal was recorded with a sampling frequency of 500 Hz. Offline analyses were conducted using Matlab 7.14 (Mathworks, Natick, MA) and Fieldtrip, a Matlab toolbox for analyzing EEG/MEG data (Oostenveld et al., 2011). Trials were only considered if the participant correctly withheld the response on a target. Artifact rejection was performed in three consecutive steps. First, muscle artifacts were removed using semi-automatic artifact rejection. Subsequently, extended infomax independent component analysis (ICA), with a weight change stop criterion of $< 10^{-7}$, was performed to identify, and reject ocular components. Finally, each trial was visually inspected for any remaining artifacts. Due to excessive ocular and movement related artifacts, 8 participants were excluded from the analysis.

Spectral analysis

Time resolved changes in spectral power were computed for each trial using Fourier analysis on 500ms sliding time windows with a step size of 50ms. This results in a frequency resolution of ~ 2 Hz. To reduce spectral leakage, a Hanning taper was applied

to each of these segments. Time-frequency windows were computed for the first 1200 ms after the presentation of the target word between 2 and 20Hz. All, transformation were applied at the individual trial level that is before averages were computed for each condition and participant. Thus, the spectral power changes reflect both induced and evoked activity in the signal. Finally, averages were normalized per participant using a common baseline between 750 and 250ms before the onset of the trial. Baseline normalization was performed as the dB change with respect to baseline:

$$signalChange(dB) = 10 * \log_{10} \left(\frac{criticalPeriod}{baseline} \right)$$

Statistical analysis

Statistical analysis was performed using a cluster-based permutation approach (Maris and Oostenveld, 2007). Cluster-based permutation reduces the number of tests in a given comparison by clustering electrodes based on a spatial adjacency and a threshold criterion ($\alpha=.05$). Subsequently, the summed t-value of each cluster is tested against a permutation distribution in which the condition labels are randomly permuted (1000 iterations). Given the a priori assumptions about the time and frequency range of the effect, based on van Ackeren and Rueschemeyer (2014), statistical analysis was performed in the theta range (4-6 Hz), and a time window of 750-1000ms after the onset of the target word.

Results

Behavioral results

Data from all 32 participants were included in the analysis of the behavioral data. To avoid movement related contamination of the EEG signal participants were asked to respond after a delay period of 2000ms. For this reason, the analysis of reaction time would not be informative. Therefore, the main behavioral analysis focused on the number of correct verifications of feature target combinations.

First, a one-sample t-test was computed on the percentage of correctly verified/falsified feature-target combinations against a test value of .5. This test revealed that overall accuracy ($M = 90\%$ $SE = 1\%$) was significantly higher than chance ($t(31)=51.79, p<.001$).

Second, a sensitivity index (d') was calculated to test whether participants could more accurately distinguish between correct and incorrect feature pairs in one of the two conditions. In signal detection theory, d' quantifies how accurately a signal (congruent

feature target combinations) is detected, taking into account the contribution from a possible response bias. Here, d' was calculated as the difference between the z -transformed hit rate (correctly identified congruent/total congruent) and the z -transformed false alarm rate (incorrectly identified incongruent/total incongruent). Subsequently, as the data was not normally distributed, a Wilcoxon sign-rank test was conducted to test for a difference between conditions. This analysis revealed that d' was overall higher for modality-specific ($M = 3.54$, $SE = .27$) versus cross-modal ($M = 2.83$, $SE = .16$) feature combinations ($z(31)=3.72$, $p<.001$). In line with previous findings, these results suggest, that participants verify features from the same modality more easily than features from different modalities.

Analysis of sustained theta power

Cluster-based permutation was performed to test for an interaction effect between veracity (congruent versus incongruent), and modality-specificity (cross-modal versus modality-specific). Specifically, the average theta power time course of the modality-specific condition was subtracted from the cross-modal condition, separately for the congruent and incongruent target feature combinations (Figure 4.2A). Subsequently, cluster-permutation was applied to test for a difference between the two difference curves. The analysis revealed one significant cluster between 750 and 1000ms after the presentation of the target word ($p=.006$), showing a fronto-central scalp topography (Figure 4.2B). To further investigate whether this effect was driven by a modality-specificity effect in the congruent condition, permutation-based paired-sample t -tests were computed to test for a difference between the cross modal versus modality-specific condition for congruent and incongruent feature-target combinations. These tests revealed a significant difference between cross-modal and modality-specific feature contexts in the congruent ($t(23)=2.098$, $p=.04$), but not the incongruent condition ($t(23)=-.364$, $p=.687$) (Figure 4.2).

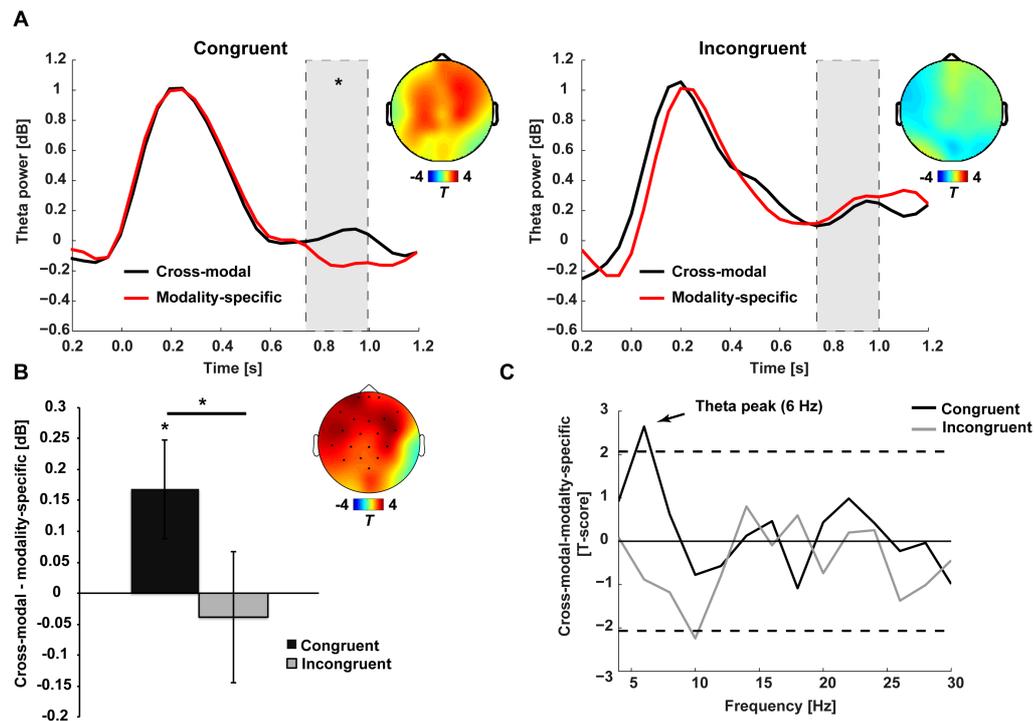


Figure 4.2. Spectral analysis of the post-stimulus period. A. The left panel depicts the theta (4-6 Hz) time course during the presentation of the target word when participants activate congruent features from different (black) or the same (red) modality. The shaded area highlights the time window of interest. The star indicates that there is a significant difference between conditions. The topography at the top illustrates the statistical contrast between the cross-modal and modality-specific condition. The right panel depicts the timecourse for the incongruent feature context. B. Bar graphs show the difference between the cross-modal and modality-specific feature context for congruent (black), and incongruent (grey) feature-target pairs. Error bars depict standard error of the mean. The stars highlight the significant interaction effect and the effect of modality-specificity for congruent feature-target combinations. The topography at the top represents the interaction effect. Black dots depict the electrode locations that were part of the significant cluster. C. Spectral decomposition in the time window of interest, and across all electrodes. A peak difference effect of modality-specificity in the theta range (6 Hz) for the congruent (black), but not the incongruent features-target pairs (grey). Dashed lines illustrate the significance threshold ($\alpha = .05$, two sided)

Discussion

In previous years, numerous studies have demonstrated that words with dominant modality-specific semantic content (color words, action verbs) engage modality-specific semantic networks in the brain (Hauk et al., 2004; Simmons et al., 2007). A recent study

(van Ackeren and Rueschemeyer, 2014) has suggested that information from these distributed semantic networks could be integrated via slow oscillatory neuronal activity in the theta. The aim of the present study was to test whether oscillatory neuronal activity in the theta range is sensitive to integrating multimodal semantic knowledge or rather switching between modalities. To this end, a dual feature verification task was used in which the same cross- and modality-specific feature pairs were either paired with a congruent or incongruent target word. As switching effects are also observed in incongruent or novel feature target combinations, a difference between cross-modal and modality-specific feature pairs should be observed in both congruent and incongruent feature-target pairs. In contrast, effects of modality-specificity should only be observed in the congruent condition if the effect reflects feature integration with the target word. Replicating previous findings by van Ackeren Rueschemeyer (2014), a more sustained theta power increase was observed in the cross modal versus modality-specific condition for congruent feature-target combinations. However, no difference was observed in the incongruent trials. These results further support the hypothesis that the delayed theta offset is indicative of cross-modal feature integration, but not modality switching.

The current study revealed enhanced theta power during the presentation of the target word when participants evaluated two features from different versus the same modality, and this difference was restricted to the congruent condition. However, it should be mentioned that there was also a relative power increase during both cross-modal and modality-specific feature pairs for the incongruent condition. There are at least two possible explanations for this finding. First, van Ackeren and Ruechemeyer (2014) showed that more sustained theta power during the cross-modal conditions is indicative of prolonged network interactions between the ATL and a widely distributed cortical network. In line with this idea, incongruent trials could be argued to elicit a more prolonged memory search. Second, enhanced theta power during incongruent trials could reflect an entirely different mechanism that is sensitive to semantic violations. Specifically, a number of studies have reported enhanced theta power during semantic violations in sentences (Hald et al., 2006; Wang et al., 2012b). The relation and possible overlap between theta modulations for error detection and memory integration or retrieval remain unclear. However, while the pattern of the congruent versus incongruent condition reveals interesting questions for future research, it does not directly affect the interpretation of the current study. Even though theta oscillation are sensitive to incongruent feature-target combinations overall, there is no added effect of modality-specificity.

The current study revealed that sustained theta power during cross-modal versus modality-specific feature verification likely reflects semantic integration between

semantic features and the target word rather than merely modality-switching. These results further substantiate the idea that slow frequency cortical oscillations are involved in language integration processes. However, similar to van Ackeren and Rueschmeyer (2014), the current study was constrained to the lower frequency range, and in particular the theta band (4-6 Hz). As suggested, by the authors local network interaction, that is, modality-specific feature integration could be reflected in higher frequency ranges (>30 Hz). Future work could address this possibility using more sensitive recording methods such as electrocorticography (ECoG) or magnetoencephalography (MEG).

Chapter 5:

Oscillatory neuronal activity reflects lexical-semantic feature integration within and across sensory modalities in distributed cortical networks

(based on: **van Ackeren MJ**, Schneider TR, Müsch K, Rueschemeyer S (in press) Oscillatory neuronal activity reflects lexical-semantic feature integration within and across sensory modalities in distributed cortical networks. *The Journal of Neuroscience*)⁴

Abstract

Research from the previous decade suggests that word meaning is partially stored in distributed modality-specific cortical networks. However, little is known about the mechanisms by which semantic content from multiple modalities is integrated into a coherent multisensory representation. Therefore we aimed to characterize differences between integration of lexical-semantic information from a single compared to two sensory modalities. We used magnetoencephalography (MEG) in humans to investigate changes in oscillatory neuronal activity while participants verified two features for a given target word (e.g., *bus*). Feature pairs consisted of either two features from the same (visual: *red-big*) or different modalities (audio-visual: *red-loud*). The results suggest that integrating modality-specific features of the target word is associated with enhanced high-frequency power (80-120 Hz), while integrating features from different modalities is associated with a sustained increase in low-frequency power (2-8 Hz). Source reconstruction revealed a peak in the anterior temporal lobe (ATL) for low- and high-frequency effects. These results suggest that integrating lexical-semantic knowledge at different cortical scales is reflected in oscillatory neuronal activity in modality-specific and multimodal association networks.

⁴ The author designed the experiment, tested the participants, analyzed the results, and wrote the manuscript. The project was supervised by Dr. Shirley-Ann Rueschemeyer, and additional support with the analysis was provided by Dr. Kathrin Müsch, and Dr. Till Schneider.

Introduction

Imaging studies have shown that processing verbal descriptions of perceptual/motoric information activate neural networks also engaged in processing this information at the sensory level. For example, colour words activate pathways that are also sensitive to chromatic contrasts (Simmons et al., 2007), action words engage areas that are involved in action planning (Hauk et al., 2004; van Dam et al., 2010), and words with acoustic associations (e.g., *telephone*) engage regions that are sensitive to meaningful sounds (Kiefer et al., 2008). More recent studies have investigated words that are associated with features from more than one modality. For example, van Dam et al. (2012) showed that words that are associated with visual and action features (e.g., *tennis ball*) activate networks from both modalities. While this is evidence that lexical-semantic knowledge is partially stored in modality-specific cortical networks, the interactions within and across these networks remain largely unexplored.

Many current models agree that modality-specific knowledge is partially stored in distinct modality-specific networks. However, where, and how this information is integrated is still debated. Some accounts postulate that information is integrated in distributed convergence zones, (Damasio et al., 2004; Barsalou, Simmons, Barbey, & Wilson, 2003), while others have argued for a single semantic hub in the ATL (Patterson et al., 2007). While there is compelling neuropsychological evidence for the existence of a semantic hub (Hodges et al., 1992), little is known about how the network dynamics within this region combine semantic content from distributed sources in the cortex. Combining multiple semantic features may be particularly relevant for identifying a specific token of a concept (e.g., a red car), and detecting relationships between concepts that are not perceptually-based.

Physiological evidence from animals and humans suggests that distributed information is integrated through synchronized oscillatory activity (Singer and Gray, 1995). Several accounts have attempted to link oscillatory activity with specific perceptual and cognitive processes, (e.g., Bastiaansen et al., 2005; Raghavachari et al., 2001; Schneider et al., 2008), however the link between power changes in a given frequency band and specific perceptual/cognitive processes remains contentious. One compelling recent account suggests that high- and low-frequency oscillations may operate at different spatial scales at the level of the cortex (Donner and Siegel, 2011; von Stein and Sarnthein, 2000). Specifically, it has been argued that oscillatory activity at low frequencies (<30 Hz) are involved in coordinating distributed neural populations whereas interactions within a neural population are reflected in high frequencies (>30 Hz). This idea converges with evidence suggesting that high-frequency oscillatory

activity can be nested in low-frequency cycles (Canolty et al., 2006). With respect to embodied accounts of lexical-semantics this is interesting, because integrating distributed information across modalities (e.g., colour-sound) may be reflected in lower frequency bands than integrating modality-specific information locally (e.g., colour-shape).

We used MEG to test whether oscillatory neuronal activity is relevant for the integration of multimodal semantic information during word comprehension. In the experimental paradigm two feature words and one target word were presented visually. Feature-words either referred to modality-specific (MS) information from the same, or cross-modal (CM) information from different modalities (Figure 5.1A). We hypothesized that integrating features of the target word from local modality-specific networks will be reflected in high-frequency oscillatory activity, while integrating features across modalities will induce a modulation in low frequencies.

Materials and Methods

Participants

Participants were twenty-two healthy, right-handed individuals (9 male) with no neurological disorder, normal or corrected to normal vision, and no known auditory impairment. The age range was 18 to 35 years ($M=24.45$, $SE=1.19$). Four participants were excluded from the analysis due to excessive ocular and movement-related artifacts. All participants were students at the University of York, and participated on a voluntary basis. Participants received either a financial reward or course credits for taking part in the experiment. All participants gave written informed consent according to the Declaration of Helsinki and were debriefed after the study. The study was approved by the Ethics Committee of the York Neuroimaging Centre.

Experimental design

Participants performed a dual property verification task. After a baseline period of fixation (1000 ms) two feature words (e.g., *red-big*) (400 ms each) and a target (e.g., *bus*) (1500 ms) were presented (Figure 5.1C). Participants were instructed to decide whether the two features can be used to describe the target. To reduce motor response related activity in the signal, participants were asked to respond only on catch trials (33%, see below).

196 unique features were extracted from previous feature rating studies (Dantzig et al., 2010; Lynott and Connell, 2009) and a rating for the purpose of the current experiment ($N=10$) in which participants rated how likely they perceived a given feature

by seeing, hearing, feeling, tasting and smelling. Dependent-sample *t*-tests confirmed that visual features were more strongly experienced by seeing as compared to hearing, feeling, tasting, and smelling (all $p < .001$). Similarly, auditory features were more strongly perceived by hearing than seeing, feeling, tasting or smelling (all $p < .001$).

300 target words were either paired with a CM, or MS feature pair. For example, *bus* was paired with visual features (*red-big*) in the MS condition, and a visual and auditory feature in the CM condition (*red-loud*). Feature pairs were rated (N=10), and matched for relatedness and predictability with respect to the target word in the CM versus MS condition ($p = .83$; $p = .123$), as well as the modality-specific visual (MS-v) versus auditory (MS-a) condition ($p = .092$; $p = .95$). In addition, features were matched between the cross-modal and modality-specific condition for word length ($p = .75$), log10 frequency ($p = .86$) (British National Corpus BNC; <http://www.kilgarriff.co.uk/bnc-readme.html>), and the proportion of gerundives to adjectives ($p < .24$).

Target words were the same in the CM and MS conditions, but were matched for word length and log10 frequency (British National Corpus BNC; <http://www.kilgarriff.co.uk/bnc-readme.html>) across the MS-v and MS-a conditions ($p = .74$; $p = .403$).

Additionally, 150 feature-target (IF) combinations (e.g., *rumbling-squeaking-cactus*) were included as non-integrable distracters. Each participant saw an experimental target word only in one condition, resulting in 100 CM, 100 MS-v, 100 MS-a, and 150 IF pairs. The presentation of experimental items in a given condition was pseudo-randomized within each participant and counterbalanced over participants.

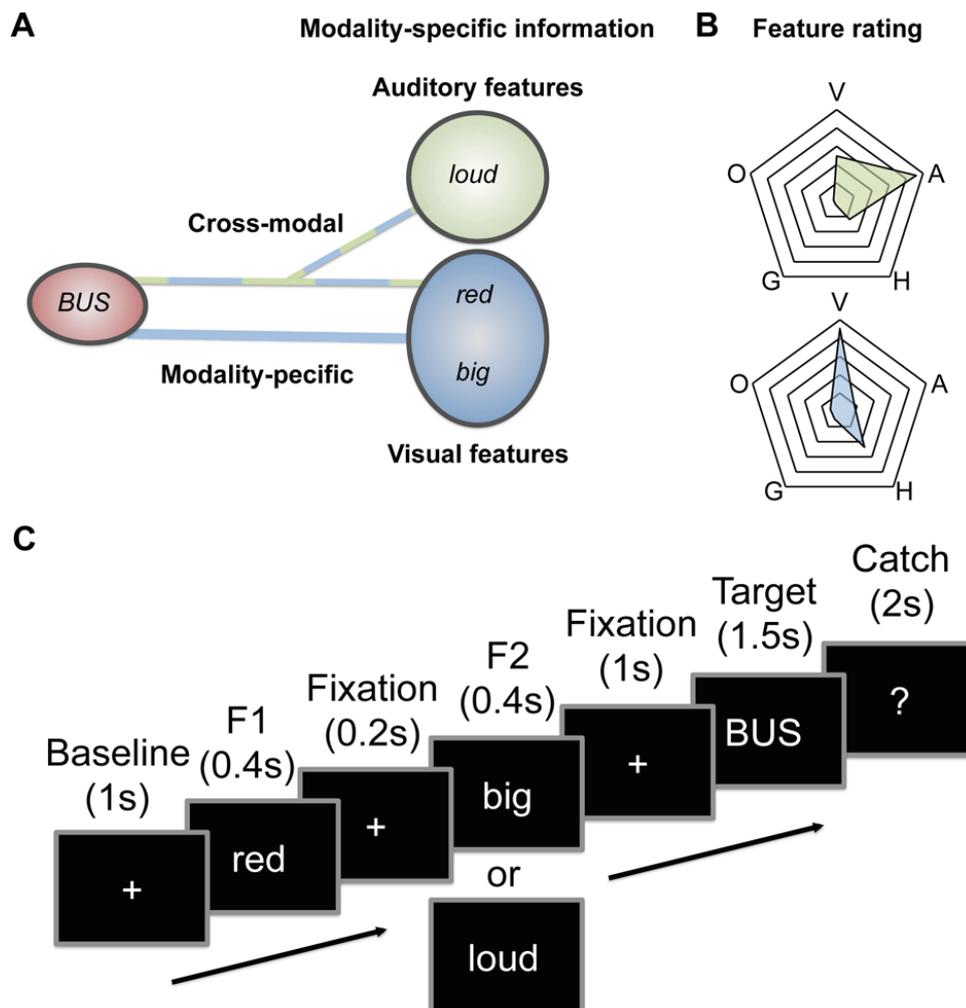


Figure 5.1. Design of the experiment. A. The word *bus* is associated with visual and auditory features. Participants are required to verify features from either the same (visual: *red-big*) or different modalities (audio-visual: *red-loud*). B Spider plots show that auditory (green) and visual features (blue) are rated as predominantly auditory (A) and visual (V) rather than haptic (H), olfactory (O), or gustatory (G). C. Features were always presented before target words.

Data acquisition

MEG data were acquired on a Magnes 3600 whole-head 248-channel magnetometer system (4-D Neuroimaging Inc., San Diego) using a sampling rate of 678.17 Hz. Head position was measured at the start and end of the experiment from five head coils (left and right preauricular points, Cz, nasion, and inion) using a Polhemus Fastrack System (Polhemus Fastrak, Colchester, VT). Horizontal and vertical eye movements as well as cardiac responses were recorded and monitored during the whole experiment.

Structural T1 weighted images (TR=8.03 ms, TE=3.07 ms, flip angle=20°, spatial resolution=1.13x1.13x1.0mm, in plane resolution = 256x256x176) were recorded on a GE 3.0 Tesla Excite HDx system (General Electric, Milwaukee, USA) using an 8-channel head coil and a 3-D fast spoiled gradient recall sequence. Co-registration of the MEG to the structural images was performed using anatomical landmarks (preauricular points, nasion), and a surface-matching technique based on individually digitized head shapes.

Preprocessing

The analysis was performed using Matlab 7.14 (MathWorks, Natick, MA) and Fieldtrip (<http://fieldtrip.fcdonders.nl/>). For subsequent analysis, the data were band-pass filtered (0.5-170 Hz, Butterworth filter, low-pass filter order = 4, high-pass filter order = 3) and resampled (400 Hz). Line-noise was suppressed by filtering the 50, 100, and 150 Hz Fourier components. Artifact rejection followed a two-step procedure. First, artifacts arising from muscle contraction, squid jumps and other non-stereotyped sources (e.g., cars, cable movement) were removed using semi-automatic artifact rejection. Second, extended infomax independent component analysis (ICA), with a weight change stop criterion of $<10^{-7}$, was applied to remove components representing ocular (eye blinks, horizontal eye movements, saccadic spikes) and cardiac signals.

Time frequency analysis

Total power was computed using a sliding window Fourier transformation for each trial with fixed time windows (500 and 200 ms) in steps of 50 ms during the first 1000 ms after the onset of the target word. In order to maximize power low (2-30 Hz) and high frequencies (30-120 Hz) were analyzed separately. For low frequencies, Fourier transformation was applied to Hanning tapered time windows of 500 ms, resulting in a frequency smoothing of ~ 2 Hz. For high frequencies, a multitaper method was applied to reduce spectral leakage (Percival & Walden, 1993); sliding time windows of 200 ms were multiplied with 3 orthogonal Slepian tapers and subjected to Fourier transformation separately. The resulting power spectra were averaged over tapers, resulting in a frequency smoothing of ± 10 Hz.

Statistical analysis

Power differences across conditions were evaluated using cluster-based randomization, which controls for multiple-comparisons by clustering neighboring samples in time, frequency, and space (Maris and Oostenveld, 2007). At the sensor-level, clustering was

performed by computing dependent-samples *t*-tests for each sensor-time point in five frequency bands of interest: theta (2-8 Hz), alpha (10-14 Hz), beta (16-30 Hz), low gamma (30-70 Hz), and high gamma (80-120 Hz), during the first 1000 ms after the onset of the target word. At the source-level, clustering was performed across space at the frequency-time window of interest. Neighboring *t*-values exceeding the cluster-level threshold (corresponding to $\alpha < 0.05$) were combined into a single cluster. Cluster-level statistics were computed by comparing the summed *t*-values of each cluster against a permutation distribution. The permutation distribution was constructed by randomly permuting the conditions (1000 iterations), and calculating the maximum cluster-statistic on each iteration.

Source reconstruction

Following the recommendation from Gross et al. (2013), statistical analysis of the contrast CM-MS was performed at the sensor level; subsequent source reconstruction was used to localize this effect.

Individual structural MRI scans were segmented and the brain compartment was used to compute a single-shell headmodel for the source analysis. This individual single-shell headmodel was used to compute the forward model (Nolte, 2003) on a regular three-dimensional grid (with 10x10x10mm spacing), containing 3000 vertices. The algorithm for the forward model, which is fully implemented in the fieldtrip toolbox is based on Nolte (2003). Specifically, the lead field for a spherical volume conductor is corrected using the superposition of basis functions, whereby the coefficients are fitted to the boundary conditions. Finally, the grid points from each individual structural image were warped to corresponding locations in an MNI template grid (International consortium for brain mapping; Montreal neurological institute (MNI), Montreal, QC, Canada) to allow for statistical analysis at the group level.

Oscillatory sources were localized at the whole-brain level using Dynamic Imaging of Coherent Sources (DICS: Gross et al., 2001; Liljeström et al., 2005). DICS is a frequency domain beamformer technique. A linear transformation is applied to the data such that activity at each location in the three dimensional grid is passed with unit gain, while the activity at all other sources is suppressed. The transformation matrix is computed from the cross-spectral density matrix (CSD), and the lead field using a constrained optimization algorithm. To avoid non-invertibility of the CSD, a regularization of 5% was introduced. The CSD matrix was computed between all MEG sensors using Hanning tapers (500-1000ms; 4-8Hz) for low frequencies, and multitapers for high frequencies (150-350ms; 7 tapers; 80-120Hz). The power in the dominant direction at each grid point was computed using single value decomposition.

Results

Behavioral analysis

Responses of one participant were not recorded, leaving 17 participants for the behavioral analysis. A one-sample *t*-test confirmed that participants were able to correctly verify trials above chance ($t(16)=7.652, p<.001, M=.77, SE=.04$). There was no significant difference between the cross-modal and the modality-specific condition ($t(16)=-1.287, p=.216, CM: M=.74, SE=.05; MS: M=.78, SE=.34$).

Low frequencies are sensitive to semantic integration at the global scale

In the low frequency range, both conditions (CM, MS) show a power increase in the theta band (2-8 Hz) as well as a decrease in the alpha (10-14 Hz) and beta band (16-30 Hz) (Figure 5.2A). Statistical comparison between conditions (Figure 5.2A, third panel) revealed enhanced theta band power (2-8 Hz; 580-1000 ms) to the target word for CM versus MS ($p=.04$, two-sided) over left lateralized magnetometers (Figure 5.2B). In other words, enhanced theta power in response to visually presented words is more sustained when participants think about a target word in the context of lexical-semantic features from different modalities.

Source reconstructions for each condition were computed at the center frequency (4-8 Hz) (Gross et al., 2001; Liljeström et al., 2005). Figure 5.2C shows the difference between conditions, expressed in z-scores. Major peaks are observed in left ATL, precuneus and around the paracentral lobule. Smaller peaks are seen in left lingual and right posterior fusiform gyrus, as well as right superior occipital and middle frontal gyrus.

To evaluate whether these effects are related to evoked activity sensitive to semantic processing (N400m), time-domain data for each condition were averaged, baseline corrected (150 ms pre-target), and converted into planar gradients (Bastiaansen and Knösche, 2000). ANOVA's with repeated measures at all sensors were computed for the N400m time window (350-550 ms), and subjected to a cluster-randomization procedure. This analysis yielded no significant clusters ($p>.7$).

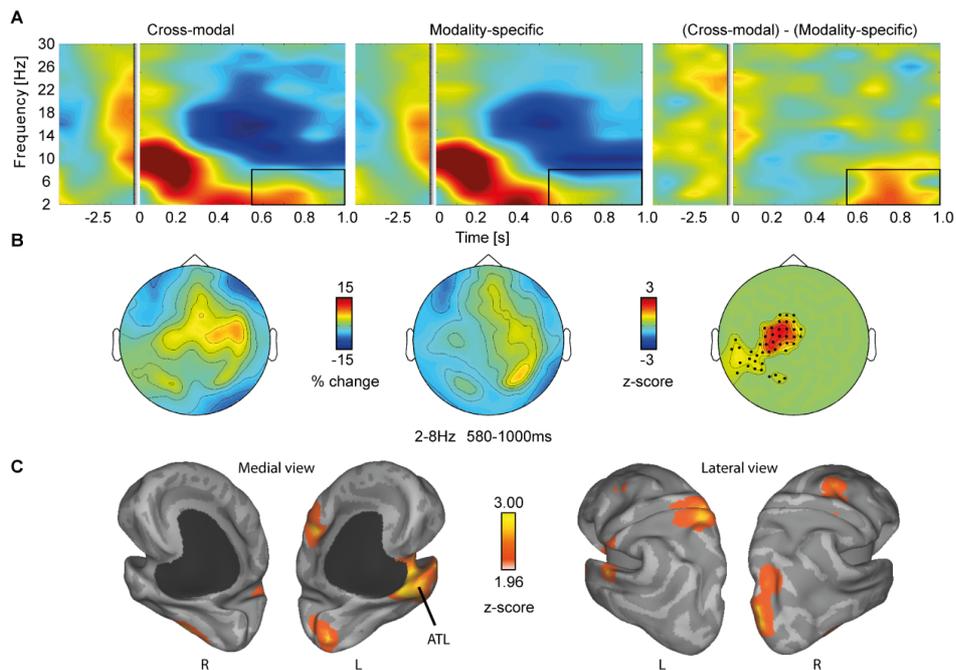


Figure 5.2. Low-frequency oscillations in sensor- and source space. A. Total power changes in the CM and MS condition relative to a 500 ms fixation baseline. Statistical differences between conditions are shown in the rightmost panel averaged over the identified cluster. The box depicts the significant time-frequency range. B. Topographies showing relative signal change across magnetometers between 2-8 Hz and 580-1000 ms. Significant channels are marked as dots in the rightmost panel. C. Source reconstruction revealed peaks in the ATL and parietal lobe.

High frequencies reflect modality-specific semantic integration

The high frequency range revealed an early increase in high-gamma power (80-120 Hz, 150-350 ms) for MS, but not CM pairs. Specifically, gamma power was enhanced when lexical-semantic features from the same modality were presented ($p=.006$, two-sided) (Figure 5.3A). The topography of the effect showed a left posterior distribution (Figure 5.3B). This suggests that integrating modality-specific features of a target word is reflected in high-frequency gamma oscillations.

Source reconstruction in the high frequency range (80-120 Hz) revealed a peak in the left ATL, as well as the medial superior frontal gyrus, mid cingulum, and left anterior cingulate cortex. Smaller peaks were observed in the right middle occipital gyrus (Figure 5.3C).

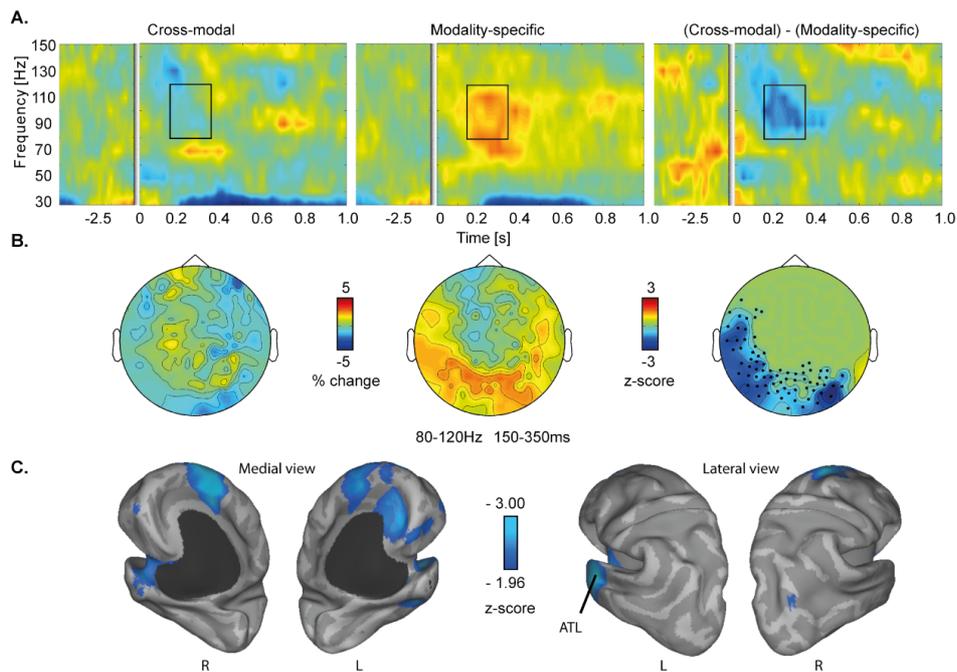


Figure 5.3. High-frequency oscillations in sensor- and source space. A. Total power changes in the CM and MS condition relative to a 500 ms fixation baseline, and the statistical difference between conditions are shown. The box depicts the significant time-frequency range. B. Topographies showing relative signal change across magnetometers between 80-120 Hz and 150-350 ms. Significant channels are marked as dots in the rightmost panel. C. Source reconstruction revealed peaks in left ATL and medial frontal lobe.

Modality-specificity to semantic features in auditory and visual cortices

Whole-brain cluster statistics in source space were performed on the two modality-specific conditions (visual and auditory) versus baseline to investigate whether enhanced gamma power reflects modality-specific network interactions. Both conditions showed enhanced gamma power in visual areas, but only the auditory condition showed a peak in left pSTS/BA22 ($p=.004$, two-sided) (Figure 5.4, first two panels). A direct comparison between conditions confirmed that gamma power in left pSTS is enhanced for the auditory, but not visual condition ($p<.002$, two-sided) (Figure 5.4, last panel). No effect in the opposite direction was found.

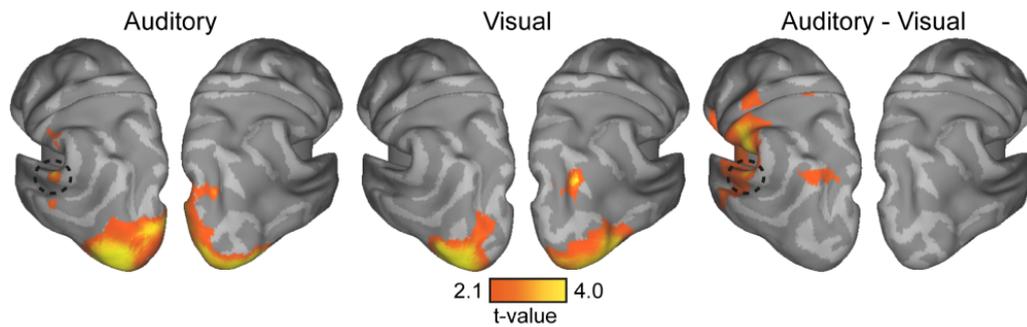


Figure 5.4. Whole brain contrasts for modality-specific auditory and visual feature contexts in the gamma range. The first two panels show the contrast between the auditory and visual condition versus baseline. The last panel depicts the contrast between the auditory and visual condition. Gamma power is enhanced in visual areas for both conditions, but only the auditory condition shows a peak in pSTS/BA22 (dashed circle). All contrasts are corrected at the cluster level ($p < .005$).

Discussion

The aim of the current experiment was to investigate a) whether lexical-semantic information is integrated in distributed modality-specific or in multi-modal hub regions and b) whether oscillatory neuronal activity contributes to the integration of semantic information from different modalities. Therefore, we aimed at characterizing differences in oscillatory neuronal activity when participants integrate semantic features of a target word (e.g., *bus*), from the same or multiple modalities. Our results suggest that integrating features from different modalities (e.g., *red-loud*) is reflected in a more sustained increase in theta power, while high gamma power is more sensitive to integrating lexical-semantic knowledge from a single modality (e.g., *red-big*). The neural generators of both effects include the ATL, a region that is proposed to be critical for semantic association and integration (Patterson, Nestor, & Rogers, 2007). Furthermore, a direct comparison between auditory and visual feature contexts revealed that gamma power is enhanced in visual areas for both conditions, while only the auditory context reveals a peak in auditory areas (pSTS)

Fast and slow components of lexical-semantic processing are reflected in the spectral and temporal profile of the signal.

Previous research using evoked fields in MEG has shown that modality-specific activation of lexical-semantic features can be detected before 200 ms (Pulvermüller, Shtyrov, & Hauk, 2009). It has been proposed that these responses reflect early parallel

processes during language comprehension. The current study revealed an early (150-350 ms) modulation in high frequency power (80-120 Hz) in response to the target word if participants were asked to access two features from the same modality. The implication of these results for neurocognitive models of language understanding is that early integration of word meaning from the same modality is reflected in a transient high frequency oscillatory response.

Furthermore, the current study identified a late theta power increase on the target word when participants verified features from different modalities. Previous research has shown that theta oscillations are sensitive to lexical-semantic content of open-class words (Bastiaansen et al., 2005). Theta is also the only known frequency band that shows a linear power increase during sentence processing, suggesting that it could be involved in ongoing integration of word meaning (Bastiaansen, van Berkum, & Hagoort, 2002). The current study revealed a power increase in the theta band when participants evaluated semantic features of the target word. We show that this response is prolonged when participants evaluate features from different modalities. A possible explanation is that integration demands increase when participants integrate information over a more widespread cortical network, that is, information from multiple modality-specific networks. This account is in line with behavioral findings showing delayed reaction times during property verification of cross-modal features (Barsalou et al., 2005)

In sum, the present study identified two processes relevant for lexical-semantic processing; an early increase in gamma power for combining similar information (i.e., MS), and a sustained increase in theta power that could reflect ongoing integration of information from distributed cortical networks.

Oscillatory neuronal activity in the ATL reflects the distribution of lexical-semantic information in the cortex.

Based on neuropsychological research in patients with temporal lobe atrophy (Hodges et al., 1992), Patterson and colleagues (2007) have argued that the ATL might be involved in combining semantic knowledge from distributed modality-specific networks. However, the physiological mechanisms of how information is integrated within, and possibly outside this region are poorly understood. The current study showed that integrating information from different modalities is reflected in sustained theta power within the ATL. As previously suggested (Raghavachari et al., 2001), theta oscillations could operate as a temporal gating mechanism in which incoming information from different networks is combined into a coherent representation.

The present data also revealed enhanced gamma power in left ATL when the target was presented with two features from the same modality. As recently demonstrated by Peelen et al. (2012), using fMRI pattern classification, the ATL is sensitive to the modality-specific content of the stimulus (motoric, visuo-spatial). This suggests that different activation patterns in the ATL could reflect inputs from distributed modality-specific networks. Indeed, recent imaging work (van Dam et al., 2012) has demonstrated that listening to words that are associated with more than one modality (visual and functional) activate multiple modality-specific cortical networks. While multiple networks respond to a single word, each network responds differently when a participant is asked to think about one modality in particular. For example, a motor network in the parietal lobe responds more strongly if the participant is asked to think about what to do with an object rather than what it looks like. The current results further demonstrate that accessing and combining modality-specific semantic information enhances gamma power in local modality-specific networks. Specifically, gamma power is enhanced in pSTS when participants are asked to access auditory features of an object. No such modulation was observed in the visual feature context. A possible reason for the lack of a semantic effect is that the sensory response to a visual word desensitizes the visual system to the more subtle semantic modulations (Pulvermüller, 2013).

Several studies have associated cross-modal perceptual matching with enhanced gamma activity (Schneider et al., 2008). These findings are not necessarily in conflict with the current framework. As Donner and Siegel (2011) point out, local gamma band modulations can also be the result of higher-order interactions. The current study extends this work showing that oscillatory dynamics in temporal association networks reflect whether one or multiple local networks participate in these interactions.

In conclusion, the current study has demonstrated that combining word meaning from a single modality is reflected in early oscillatory activity in the gamma band, originating in sensory cortices and left ATL, respectively. Precisely, modality-specific networks in the auditory cortex were more sensitive to auditory than visual features. In contrast, integrating features from multiple modalities induced a more sustained oscillatory response in the theta band that was partially localized to ventral networks in the ATL. Taken together, these results represent a mechanistic framework for lexical-semantic processing. At the physiological level, accessing knowledge from a single or multiple semantic networks is reflected in oscillatory activity at different frequencies. At the cognitive level, the current data suggest two processes that operate in parallel, but at a different temporal resolution; a fast process for combining similar information early

on, and a slow process that could be involved in integrating distributed semantic information into a coherent object representation.

Chapter 6:

Slow cortical oscillations connect left anterior temporal lobe to modality-specific cortical networks during semantic feature retrieval and integration

Abstract

Language understanding rests on the ability to retrieve and integrate rich semantic information from long-term memory. Current neurocognitive models argue that this knowledge is represented in distributed modality-specific cortical networks. For example, visual information is represented in ventral visual networks, while auditory information is represented in superior temporal auditory networks. Some theories posit that these networks are connected via a central hub in the anterior temporal lobe (ATL), which allows us to integrate features from multiple modalities and find relations between objects that are not perceptually-based. The aim of the current study is to investigate the neurophysiological mechanisms of how the ATL-hub and modality-specific networks interact during semantic feature retrieval and integration. Magnetoencephalography (MEG) was used to record oscillatory neuronal activity, as participants performed a dual feature verification task. Participants saw two features (e.g., *red* and *big*) followed by a target word (*bus*), and were asked to indicate whether the feature pair is consistent with the target word. Feature pairs consisted of two visual features, two auditory features, or one auditory and one visual feature. Using the phase-locking value (PLV) in low frequency theta oscillations as a proxy for long-range connectivity, the present study demonstrates that the connection weights between ATL and modality-specific networks are modulated consistently with the type of feature the participant is asked to retrieve. That is, thinking about visual features of the target word selectively enhances the connection between ATL and ventral visual networks (inferior lateral occipital), while thinking about auditory features selectively enhances the connection between ATL and auditory networks (superior temporal). These results corroborate and extend earlier findings, suggesting that the ATL directly exchanges information with distributed modality-specific networks in a context-specific way. Furthermore, these interactions are reflected slow cortical oscillations in the theta range.

Introduction

Neurocognitive models of semantic cognition agree that accessing perceptual semantic information both verbally, and non-verbally engages distributed modality-specific cortical networks (Barsalou et al., 2003; Binder et al., 2009; Vigliocco et al., 2009; Pulvermüller and Fadiga, 2010; Binder and Desai, 2011). In recent years numerous studies have provided support for this idea using a range of behavioral and neuroimaging methods (Bonner & Grossman, 2012; Goldberg, Perfetti, & Schneider, 2006; Hauk et al., 2004; Kiefer et al., 2008; Martin & Chao, 2001; Simmons et al., 2007). For example, words describing visual information such as color and shape engage ventral visual networks in the inferior occipital and temporal lobe (Pulvermüller and Hauk, 2006; Simmons et al., 2007; van Dam et al., 2012), while words with dominant auditory associations engage auditory networks in the superior and middle temporal lobe (Bonner & Grossman, 2012; Goldberg, Perfetti, & Schneider, 2006; Kiefer et al., 2008). According to the hub-and-spokes model, an influential theoretical framework in semantic cognition, these distributed cortical networks converge in a single putative hub in the anterior temporal lobe (ATL). However, the neurophysiological mechanisms of how the ATL-hub interacts with modality-specific spokes are still poorly understood. The current study directly addresses this issue by testing whether direct connections between modality-specific networks and the ATL, are modulated by semantic word content.

Connectionist models have demonstrated that distributed modality-specific networks alone are insufficient to explain the complex taxonomic organization of the semantic system in humans (McClelland, Rogers, & others, 2003; Plaut, 2002). Thus, current theories postulate that information converges in multiple convergence zones (Binder & Desai, 2011; Damasio, 1989; Pulvermüller, 2013) or a single supramodal hub (Patterson et al., 2007). According to the *hub-and spokes model* information from multiple modalities converges in a single hub that is thought to be represented in the ATL. The model has received substantial support from patients with semantic dementia (SD) who show severe atrophy in bilateral ATL (Hodges et al., 1992). As the disease progresses patients categorize objects based on simple perceptual features rather than abstract category membership. This can result in undergeneralizations for atypical members of a category, or overgeneralizations for perceptually similar non-members of a category (Lambon Ralph et al., 2010).

More recent studies have used fMRI multi-voxel pattern analysis (MVPA) in healthy individuals to investigate directly what information is represented in ATL, and modality-specific networks. For example, Peelen and Caramazza (2012) presented

images of objects that differed in terms of how they are used (action) or where they can be found (visuo-spatial). The authors found that classification accuracy for the conceptual distinction was above chance in the anterior temporal lobe, while posterior regions were sensitive to the perceptual, or pixel-wise similarities between stimuli. Using a different approach, Coutanche and Thompson-Schill (2014) presented a verbal cue for a visual object (one out of four different types of food) that, after a delay period, emerged from Gaussian visual noise. Importantly, objects either shared the same shape or color. The authors found enhanced classification accuracy for shape (round versus elongated) in lateral occipital cortex (LOC), while color was represented in visual area V4. In contrast, classification of food identity only reached significance in left ATL. In addition, the authors showed that classification success in ATL was related to classification success in both LOC and V4. This is interesting because it shows a direct relationship between the representational content in ATL and modality-specific cortical networks.

Neuropsychological and functional imaging studies have provided important insights into the different components of the semantic system, their representational content, and mutual dependencies, yet the neurophysiological processes of how modality-specific spokes and the putative hub in ATL communicate are still poorly understood. It has been hypothesized recently that semantic processing at different cortical scales may be reflected in oscillatory neuronal activity at different frequencies (van Ackeren et al., in press; van Ackeren and Rueschemeyer, 2014). For example, van Ackeren and colleagues (in press) used MEG to study changes in oscillatory neuronal activity when participants verified two features of visually presented words. Feature pairs consisted of features from the same (auditory or visual) or different modalities (audio-visual). The authors report that low frequency theta power in left ATL is enhanced when participants access distributed (cross-modal) information, while high frequency gamma power is enhanced in left ATL and modality-specific networks when information is accessed locally (within modalities). In other words, the way semantic information is distributed in the cortex (local versus distributed) is reflected in oscillatory dynamics at different frequencies (high gamma versus low theta). In a different study van Ackeren and Rueschemeyer (2014) found that the theta power increase is directly associated with enhanced theta phase coupling between left ATL and a widespread cortical network. Phase coupling is considered a proxy for long-range connectivity between distant areas. Based on these results, the authors suggest that slow cortical oscillations could be involved in integrating multimodal semantic knowledge into coherent object representations by linking modality-specific semantic networks to a hub in left ATL.

While van Ackeren and Rueschemeyer (2014) demonstrate that slow cortical oscillations are involved in long-range network interactions, it is not clear whether these distributed patterns of activation are indeed the result of modality-specific access or rather a reflection of the general processing disadvantage for integrating cross-modal semantic features. Specifically, if slow cortical oscillations are involved in linking left ATL to modality-specific regions, one would expect that the connection weights change when participants access features from a single, or multiple modalities (Kiefer et al., 2008; van Dam et al., 2012). This hypothesis was directly tested in the current study. Using the phase locking value (PLV) as a proxy for long-range connectivity (Lachaux et al., 1999), we tested whether long-range interactions between the ATL-hub and modality-specific spokes in visual (inferior occipital) and auditory (superior/middle temporal) networks is modulated as participants retrieve and integrate different semantic features of a target word. The analysis was conducted on the dataset from van Ackeren et al (in press), which is also discussed in Chapter 5. In line with predictions from the hub –and-spokes model, we hypothesize that the connection weights between left ATL and a left auditory network (left STG) are enhanced when participant think about auditory, rather than visual features of a target word. In contrast, the connection weights between left ATL and a ventral visual network (iOC) should be enhanced for visual versus auditory features. Lastly, we predict that integrating two features denoting auditory and visual information should enhance both ATL-auditory, and ATL-visual connection weights.

Materials and Methods

Participants

Participants were twenty-two healthy, right-handed individuals (9 male) with no neurological disorder, normal or corrected to normal vision, and no known auditory impairment. The age range was 18 to 35 years ($M=24.45$, $SE=1.19$). Four participants were excluded from the analysis due to excessive ocular and movement-related artifacts. All participants were students at the University of York, and participated on a voluntary basis. Participants received either a financial reward or course credits for taking part in the experiment. All participants gave written informed consent according to the Declaration of Helsinki and were debriefed after the study. The study was approved by the Ethics Committee of the York Neuroimaging Centre.

Experimental Design

The design of the study is illustrated in Figure 6.1. Participants were asked to verify whether two feature words are consistent with a target word. The trial started with a period of fixation (1000ms), followed by two feature words (400ms each) that were separated in time by another brief fixation period (200ms). Finally, the target word was presented (1500ms) after another fixation period (1000ms). Participants were asked to respond on catch trials (33%) only to reduce motor related contamination of the signal.

Target words consistent of 300 nouns that were either paired with a cross-modal or modality-specific feature pair from a pool of 196 unique auditory and visual features. Thus the same target words were presented in the cross-modal and modality-specific condition. For the modality-specific auditory and visual condition, target words were matched for word length and log10 frequency (British National Corpus BNC; <http://www.kilgarriff.co.uk/bnc-readme.html>) ($p=.74$; $p=.403$). Features rated as predominantly visual or auditory were selected based on multiple databases (for details of the matching procedure see van Ackeren, Schneider, Müsch, & Rueschemeyer, in press), and matched for relatedness and predictability with respect to the target word in the cross-modal versus modality-specific condition ($p=.83$; $p=.123$), as well as the modality-specific visual versus auditory condition ($p=.092$; $p=.95$). In addition to 150 incongruent feature target pairs were included in the experiment as distracters. Thus, taken together each participant saw 100 cross-modal trials, 100 visual, 100 auditory, and 150 incongruent feature pairs.

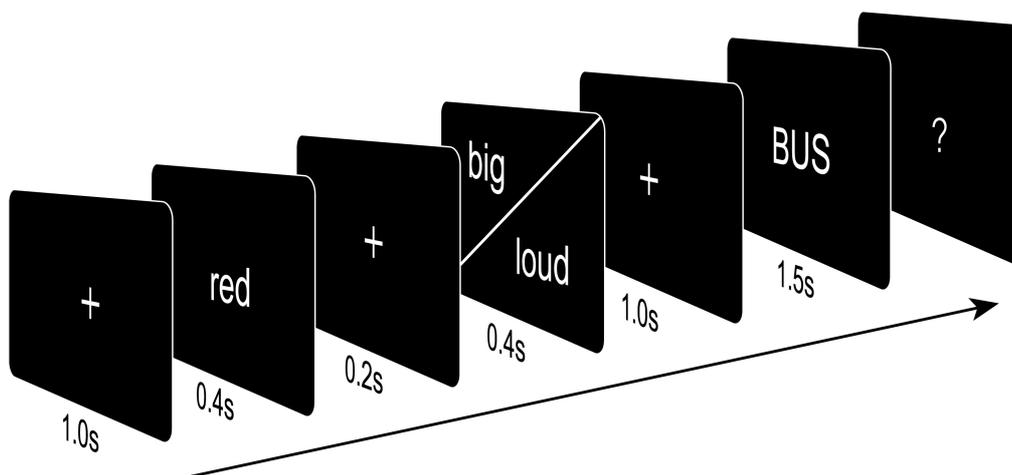


Figure 6.1. Design of the dual property verification task. After a period of fixation, two feature words are presented followed by the target word. Participants are asked to respond on catch trials only.

Data acquisition

MEG data were acquired on a Magnes 3600 whole-head 248-channel magnetometer system (4-D Neuroimaging Inc., San Diego) using a sampling rate of 678.17 Hz. Head position was measured at the start and end of the experiment from five head coils (left and right preauricular points, Cz, nasion, andinion) using a Polhemus Fastrack System (Polhemus Fastrak, Colchester, VT). Horizontal and vertical eye movements as well as cardiac responses were recorded and monitored during the whole experiment.

Structural T1 weighted images (TR=8.03 ms, TE=3.07 ms, flip angle=20°, spatial resolution=1.13x1.13x1.0mm, in plane resolution = 256x256x176) were recorded on a GE 3.0 Tesla Excite HDx system (General Electric, Milwaukee, USA) using an 8-channel head coil and a 3-D fast spoiled gradient recall sequence. Co-registration of the MEG to the structural images was performed using anatomical landmarks (preauricular points, nasion), and a surface-matching technique based on individually digitized head shapes.

Data analysis and statistical testing

The analysis was performed using Matlab 7.14 (MathWorks, Natick, MA) and Fieldtrip (<http://fieldtrip.fcdonders.nl/>). For subsequent analysis, the data were band-pass filtered (0.5-170 Hz, Butterworth filter, low-pass filter order = 4, high-pass filter order = 3) and resampled (400 Hz). Line-noise was suppressed by filtering the 50, 100, and 150 Hz Fourier components. Artifact rejection followed a two-step procedure. First, artifacts arising from muscle contraction, squid jumps and other non-stereotyped sources (e.g., cars, cable movement) were removed using semi-automatic artifact rejection. Second, extended infomax independent component analysis (ICA), with a weight change stop criterion of $<10^{-7}$, was applied to remove components representing ocular (eye blinks, horizontal eye movements, saccadic spikes) and cardiac signals.

The current source space analysis focuses on the long-range connectivity reflected in oscillatory dynamics in the theta range (4-8 Hz) between the left ATL as a putative hub region (Figure 6.2 bottom) and modality-specific visual (left inferior occipital, red contours in Figure 6.2), and auditory (left superior temporal, blue contours in Figure 6.2) regions as putative spokes regions.

Complex Fourier transformations were performed around 5 ± 2 Hz for a time window between 0 and 800ms after the onset of the target word. Prior to transformation, a Hanning taper was applied in the time domain to reduce spectral leakage. For subsequent source analysis, individual trial Fourier spectra were converted to a cross-spectral density (CSD) matrix comprising all MEG sensor combinations.

Source space analysis was performed using a linear beamforming method in the frequency domain (Gross et al., 2001; Liljeström et al., 2005). Individual brain compartments were extracted from structural MRI images and used to create a single-shell headmodel on a regular three-dimensional grid (10^3 mm spacing) for source analysis (Nolte, 2003). To allow source space statistics at the group level, individual headmodels were aligned to standard mni space (International consortium for brain mapping; Montreal neurological institute (MNI), Montreal, QC, Canada). Subsequently, spatial filters were computed for each point in the source grid using the leadfield and the CSD matrix across all conditions.

For the functional connectivity analysis, a seed region was chosen in left ATL, at the grid point which was most sensitive to the power difference between the cross-modal and modality-specific condition in the theta band [mni: -20 -10 -30] (van Ackeren et al., in press). The exact anatomical location of this region is the left parahippocampal gyrus. As connectivity measures are sensitive to differences in trial numbers, a random selection procedure was used to extract the same number of trials from all conditions. Subsequently, the spatial filter coefficients, common to all conditions, were used to project Fourier spectra for each individual trial to each grid location in the source model. Individual Fourier spectra were used to compute a CSD matrix between all grid locations inside the brain compartment for each condition and participant. Finally CSD matrices were used to compute the phase locking value (PLV), which can be considered a proxy for long-range connectivity (Lachaux et al., 1999). Here, PLV is used to quantify to what extent the phase (ϕ) of the theta rhythm at all grid points in the brain is locked to the phase in the seed region (left ATL). In other words, a high PLV indicates a consistent time lag between the seed region and any other location in the brain, while a low PLV suggests no consistent relationship between the two locations in the brain. Finally, a stabilizing Fischer-z-transform was applied using the inverse hyperbolic tangent (\tanh^{-1}).

Statistical analysis in source space was performed on functional connectivity maps connecting left ATL to the rest of the network, using a cluster-based randomization technique. Cluster randomization reduces the number of individual tests, by clustering neighboring locations in the brain that exceed an arbitrary threshold ($\alpha=.05$). Subsequently, the size of the cluster is compared against the distribution of largest clusters that was derived from 1000 permutations of the condition labels. Given the strong a priori anatomical predictions for the modality-specific visual and auditory conditions, and to maximize sensitivity of the analysis, the search space of the clustering

algorithm was reduced to left auditory cortex, and inferior visual areas using masks from the BrainMap project (Fox et al., 1994).

Results

Behavioral analysis

A one sample t-test on the responses from 17 participants showed that participants were able to perform the task above chance. ($t(16)=7.652, p<.001, M=.77, SE=.04$). There was no significant difference between the cross-modal and the modality-specific ($t(16)=-1.287, p=.216, CM: M=.74, SE=.05; MS: M=.78, SE=.03$), or the two modality-specific conditions ($t(16)=1.0, p=.332$).

Seed-based functional connectivity analysis in left ATL

Cluster-based permutation was performed on a reduced search space comprising ventral visual and auditory areas to test for a difference in functional connectivity to left ATL between target words that were presented in the context of two visual, or two auditory features (Figure 6.3). The analysis revealed a significant negative cluster in the auditory mask (cluster-size: $50\text{mm}^3; p=.036$, two-sided), suggesting that auditory regions are more strongly connected to left ATL if participants are thinking about auditory versus visual features of the target word. In addition, a significant positive cluster was found in the ventral visual mask (cluster-size: $60\text{mm}^3; p=.028$, two-sided), revealing enhanced functional connectivity between left ATL and ventral visual areas when participants think about visual versus auditory features of an object. Taken together, these results suggest that functional connections between left ATL and modality-specific ventral visual, and auditory networks are modulated selectively when participants retrieve visual or auditory word meaning respectively.

If word meaning from a single modality selectively modulates functional connectivity between left ATL and modality-specific visual or auditory networks, accessing features from both modalities should enhance both pathways. To further test this hypothesis permutation-based t-tests with 1000 iterations were conducted comparing each modality-specific condition to the cross-modal condition in auditory and visual regions of interest. A permutation test was chosen because the data were not normally distributed. In the auditory ROI, there was a marginally significant difference between the visual and the cross-modal condition ($t(17)=-1.92, p=.068$), but no difference between the auditory and the cross-modal condition ($t(17)=.58, p=.589$). In ventral visual areas, this pattern was reversed. There was a marginally significant difference between the auditory and the cross-modal condition ($t(17)=-1.88, p=.054$),

but no difference between the visual and cross-modal condition ($t(17)=.60, p=.543$). Although these results are not considered statistically significant, we observe a trend, suggesting that accessing features from multiple modalities show a similar connectivity pattern as accessing features from a single modality in the congruent region (e.g. accessing visual features enhances functional coupling between ventral visual networks and left ATL).

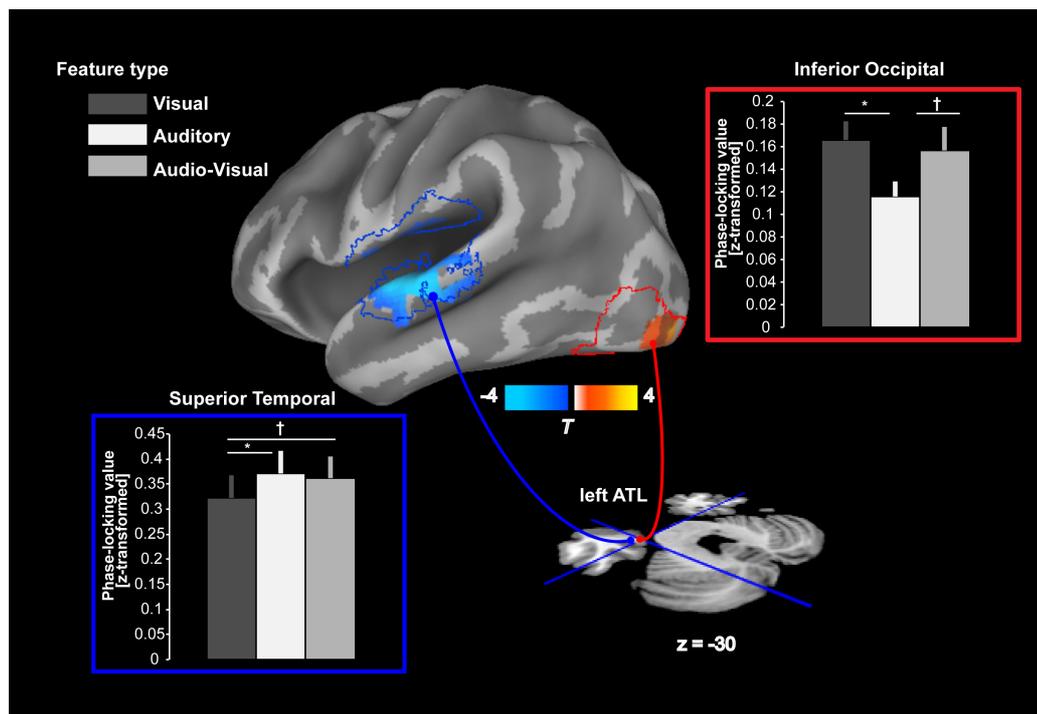


Figure 6.2. Seed-based connectivity analysis between left ATL, ventral visual (red), and temporal auditory (blue) regions. The activation map illustrates the t-contrast between modality-specific visual (e.g., green-round-APPLE) versus auditory (noisy-thumping-DRYER) feature integration (* $p < .05$; † $p < .08$).

Discussion

Despite compelling evidence that left ATL is sensitive to specific semantic categories (Lambon Ralph et al., 2010; Pobric et al., 2010a; Coutanche and Thompson-Schill, 2014) and widely connected to a distributed cortical network both structurally (Catani and Thiebaut de Schotten, 2008) and functionally (Hipp et al., 2012; van Ackeren and Rueschemeyer, 2014), there is little evidence for a direct contextual modulation of the connection weights between left ATL and modality-specific networks. The present study

directly addressed this problem. Participants were probed to think about auditory, visual or audio-visual features of a target noun while electromagnetic activity was recorded from the scalp using MEG. Seeded functional connectivity in the theta band was computed between left ATL and modality-specific visual (inferior occipital), and auditory (superior/middle temporal) networks. Based on the hub-and-spokes model and previous work (van Ackeren and Rueschemeyer, 2014), we predicted that long-range connections between left ATL and modality-specific networks should be selectively enhanced when participants access visual versus auditory word content. The current results confirmed this prediction. Specifically, the current study demonstrated that a) functional connectivity between left ATL and distributed semantic networks is mediated by (modality-specific) word content, and b) these long-range connectivity patterns are supported by slow cortical oscillations in the theta band.

Distributed semantic networks are linked via left ATL

Current neurocognitive models of semantic cognition agree that information is partially stored in distributed modality-specific networks (Binder & Desai, 2011; Damasio, 1989; Patterson et al., 2007; Tranel et al., 1997). According to the hub-and-spokes model, these distributed modality-specific networks are linked via a central hub in the ATL (Patterson et al., 2007). However, the neurophysiological mechanisms of how ATL and modality-specific networks communicate are currently not well understood. The present study demonstrates that accessing modality-specific semantic knowledge directly modulates the connection weights between left ATL and modality-specific networks. These results strongly support the hub-and-spokes model of semantic cognition, which postulates that concrete perceptual knowledge from distributed semantic networks converges in the anterior part of the temporal lobe.

The focus of the current study was on accessing and combining perceptual features from one or multiple modalities, which in previous studies has been strongly associated with the left ATL (Chiou et al., 2013; Coutanche and Thompson-Schill, 2014; van Ackeren and Rueschemeyer, 2014). In previous years, other accounts have argued that information is integrated in multiple distributed convergence zones (Barsalou et al., 2003; Binder et al., 2009; Binder & Desai, 2011; Damasio, Tranel, Grabowski, Adolphs, & Damasio, 2004; Hagoort, 2005). These convergence zones are thought to be sensitive to specific feature combinations, and likely integrate information over a more local cortical scale than what is predicted about left ATL. While the design of the current study was not optimized to test any additional predictions about these convergence zones, oscillatory dynamics could provide useful insights about the neurophysiological mechanisms in these regions in the future

The present study also corroborates research on patients with semantic dementia who show a bias towards categorizing objects according to perceptual rather than conceptual similarity (Lambon Ralph et al., 2010). Specifically, damage to the intermediate layer in the ATL that connects modality-specific regions could prevent any cross talk between multiple semantic networks. This would explain SD patients' reliance on simple perceptual features in categorizing concepts. Furthermore, the current results complement earlier findings by Coutanche and Thompson-Schill (2014) who showed that encoding success of individual items in left ATL depends on encoding success of perceptual features (shape and color) in posterior regions (inferior occipital/temporal lobe). Specifically, these dependencies could be established via direct long-range functional connectivity between hub and spokes as described in the present study.

Finally, there was partial evidence that accessing features from different modalities simultaneously modulates ATL-visual, as well as ATL-auditory connections. This result is in line with recent evidence suggesting that words that are associated with motoric and visual information engage both modality-specific systems (van Dam et al., 2012). Hoenig and colleagues (2008) further demonstrated that the activity in visual and motoric networks changes depending on which aspect of the object the participant is currently thinking about. In the context of the present study, one might predict that these patterns of activation could be the result of a modulation in the connection strength to left ATL, which is an interesting empirical question for future research.

Taken together, the current findings provide an important contribution to the literature on semantic cognition, showing that the connection strength between left ATL and modality-specific cortical networks is indeed sensitive to which aspects of word meaning are accessed or integrated. Thus far, these functional interactions were mostly assumed or investigated by focusing on ATL and modality-specific networks in isolation.

Slow oscillations support long-range interactions during semantic retrieval

While many theories assert that the ATL operates as a central hub region during semantic cognition, little is known about the underlying neurophysiological processes that allow this region to communicate with a distributed cortical network. The present study showed that left ATL communicates with modality-specific regions in a task-dependent way through slow cortical oscillations in the theta band. In addition, the current findings fall into a larger pool of previous research in animals and humans, suggesting that theta phase synchronization between temporal lobes and more widespread cortical structures is involved in memory processes at multiple levels such as working memory maintenance, and long-term memory encoding (Buzsáki, 2005;

Klimesch et al., 2001; Lisman, 2010; O'Keefe & Recce, 1993; Raghavachari et al., 2001).

The relationship between temporal lobes and theta oscillations has also been demonstrated in a recent MEG investigation of resting state networks (Hipp et al., 2012). Specifically, the authors showed that temporal lobes show hub-like connectivity patterns (e.g., high centrality), particularly at the theta range. In contrast, parietal association cortices showed hub-like properties at higher frequency bands. The current study complements these earlier results, showing functional changes in these connection patterns in the context of semantic retrieval.

Possible confounds resulting from field spread

In previous work by van Ackeren and Rueschemeyer (2014), the authors used the imaginary part of coherency to study long-range network interactions between left ATL and a widespread cortical network. The use of this measure is particularly advantageous for EEG data, as it is not sensitive to spurious connectivity patterns resulting from field spread (Nolte et al., 2004). However, imaginary coherency is a conservative measure, which trades in reliability for reduced sensitivity (Gross et al., 2012). In the present study, phase locking (PLV) of oscillatory neuronal activity between two locations in the cortex was used as a proxy for long-range connectivity (Lachaux et al., 1999). PLV is more likely to detect functional connectivity patterns at the price of also being more sensitive to field spread. There were several reasons for choosing PLV over imaginary coherency. First, MEG has a more focal source solution than EEG, which effectively reduces field spread. The main reason for this is the fact that EEG is very sensitive to secondary currents, or volume conduction, whereas MEG is mostly sensitive to primary currents (Baillet, 2001). Additional reasons for a more accurate source model estimation in MEG include the fact that MEG usually uses more sensors, and is acquired in a magnetically shielded room that should reduce the effects of particularly high frequency noise sources. Second, field spread is particularly strong around the seed region. However, statistical contrasts in the present study were computed by comparing using two conditions (with equalized trial numbers) directly using. Field spread should be similar in the two conditions, and therefore be cancelled out. Lastly, the problem of field spread was reduced as the connectivity analysis was performed after the source-reconstruction step.

Conclusion

The current study reanalyzed the MEG dataset by van Ackeren et al. (in press), to test whether accessing modality-specific word content (auditory, visual, or audiovisual features) of a target noun is reflected in the functional connectivity patterns between supramodal left ATL and modality-specific cortical networks in inferior occipital (visual) and superior temporal lobes (auditory). The results showed enhanced functional connectivity between left ATL and inferior occipital networks when participants thought about visual features, while accessing auditory features was reflected in enhanced connectivity between left ATL and superior temporal regions. Furthermore, there was a trend in the data suggesting that accessing word content from both modalities enhances connection weights between left ATL and modality-specific auditory and visual regions equally. These results support neurocognitive models of semantic cognition, which postulate that distributed perceptual information is linked via ATL (Patterson et al., 2007). Furthermore, the current study corroborates neurophysiological evidence suggesting that slow cortical oscillations in the theta band are involved in linking distributed semantic content (van Ackeren and Rueschemeyer, 2014).

Chapter 7:

Towards a neurophysiological framework for semantic feature selection and integration

The aim of the current thesis was to investigate the neurophysiological processes underlying selective retrieval and integration of word meaning using oscillatory neuronal dynamics. Oscillatory dynamics have been studied in the context of lexical-semantic retrieval before (Bastiaansen, van der Linden, Ter Keurs, Dijkstra, & Hagoort, 2005; Bastiaansen, Oostenveld, Jensen, & Hagoort, 2008; Hagoort, Hald, Bastiaansen, & Petersson, 2004; Pulvermüller, Lutzenberger, & Preissl, 1999), however, there is considerable variability in these results and it is not clear how to interpret them with respect to existing models of semantic cognition. A possible explanation for the lack of consistency is that hypotheses in psycholinguistic research are formulated at a different level of processing than the level at which oscillatory dynamics operate (Marr, 1982). While psycholinguistic research aims to identify specific cognitive computations, and how they are carried out, oscillatory dynamics could reflect how these computations are realized in an adaptive biological system. I discuss this view in Chapter 1 ('A dynamic view on cognition').

The approach in the present thesis differs fundamentally from previous work, as the goal was not to relate oscillatory dynamics to cognitive processes directly, but to different classes of network interactions that are inherent to these processes. This strategy was motivated by current frameworks on cortical oscillations in humans (Donner & Siegel, 2011; Engel & Fries, 2010; Hanslmayr et al., 2012; Jensen & Mazaheri, 2010; Singer & Gray, 1995), which argue that oscillatory dynamics at different frequencies reflect more generic neurophysiological mechanisms, such as inhibition (Jensen & Mazaheri, 2010), local, and long-range communication (Donner and Siegel, 2011), and the amount of information encoded in the system (Hanslmayr et al., 2012), which jointly affect multiple cognitive systems. The empirical work presented in the current thesis has implications for two components of word level processing. These are a) the selection of perceptual semantic features, and b) the integration of word meaning in local and distributed cortical networks. The next section provides a brief summary of the results from all five empirical chapters. Subsequently, the findings are discussed in the broader context of current research on word retrieval. The discussion concludes with the limitations of the work presented as well as recommendations for future studies.

Summary of the main findings

Chapter 2

Recent studies have demonstrated that the neural response in modality-specific semantic networks during word retrieval is highly context dependent (Hoenig et al., 2008; Raposo et al., 2009; Willems and Casasanto, 2011; van Ackeren et al., 2012; van Dam et al., 2012). For example, the motor system responds more strongly to a word like *hammer*, when participants think about how to use the hammer versus what it looks like (van Dam et al., 2012). However, the neural mechanisms that allow us to selectively retrieve specific perceptual features of a word are still poorly understood.

The experiment described in Chapter 2 investigates this problem under the assumption that perceptual complexity, or the number of features in long-term memory, differs between different modalities. For example, there are more ways to describe what most objects look like (visual), than what they feel like (haptic). Participants were presented with concrete nouns and asked to provide two features of the target word describing either what it looks like (visual), sounds like (auditory), or feels like (haptic), while EEG was recorded from the scalp. Features in each modality were used to compute the feature entropy as a proxy for perceptual complexity. Entropy is a concept from mathematical information theory (Shannon and Weaver, 1949), which can be used to quantify the information richness or predictability of an event.

The results reveal an early difference (~220ms) in the evoked responses to the target word when participants retrieve features from different modalities, suggesting that feature selection starts early during word processing. This finding is in line with recent research from other groups (Pulvermüller et al., 2005b, 2009; Kounios et al., 2009; Rabovsky et al., 2012). However, the main finding is that retrieving more complex information is inversely related to oscillatory power in the beta band. In other words, retrieving information from perceptually complex modalities (e.g., visual) induces a stronger reduction in oscillatory power than retrieving information from less complex modalities (e.g., haptic). The study is in line with a recent framework suggesting that a reduction in oscillatory power could signal an increase in the amount of information represented in the system at a given moment in time (Hanslmayr et al., 2012).

Chapter 3

Current neurocognitive models of language argue that word meaning is at least partially represented in a widely distributed cortical network (Patterson et al., 2007; Vigliocco et al., 2009; Pulvermüller and Fadiga, 2010; Binder and Desai, 2011). However, there is

currently a debate in the field as to where, and possibly how, this distributed information converges and is combined into a coherent representation. Chapter 3 reports a behavioral and an EEG experiment, which aim to address these questions. Both experiments use a dual property verification paradigm: Participants were asked to indicate whether two feature words are consistent with a given target word (*WHISTLE*). Feature pairs were either from the same (*tiny, silver*) or different modalities (*tiny, loud*).

The behavioral experiment demonstrates that integrating features across modalities incurs a robust processing cost as compared to integrating features from the same modality. The EEG experiment shows that integrating cross-modal information induces a more sustained increase in low frequency theta power. Source reconstruction of the effect reveals that the effect peaks in the left ATL, a region that is thought to link semantic content from distributed cortical networks (Patterson et al., 2007; Pobric et al., 2010a, 2010b; Chiou et al., 2013; Coutanche and Thompson-Schill, 2014). The likely function of theta oscillations in linking distributed networks was confirmed in a seed-based connectivity analysis; theta oscillations in the left ATL show a stable phase relationship with theta oscillations in a distributed cortical network. In line with the power analysis, these functional networks are more sustained when participants integrate cross-modal information.

Chapter 4

The EEG experiment in Chapter 3 demonstrates that theta oscillations are enhanced during cross-modal feature integration, likely reflecting more sustained network interactions between left ATL and a widespread cortical network. Yet, it is unclear whether these results reflect the integration of cross-modal information or rather modality-switching costs (Pecher et al., 2003). To address this problem, Chapter 4 reports an EEG experiment using dual property verification where the same cross-modal or modality-specific feature pairs were either presented with a congruent, or incongruent target word. If the effect in Chapter 3 is sensitive to integrating rather than switching between modalities, a relative difference between cross-modal and modality-specific feature pairs is expected only for congruent, but not incongruent condition. Indeed, the results reveal an interaction effect between modality-specificity and congruency, confirming that theta oscillations are sensitive to the integration of multimodal, or distributed, semantic content.

Chapter 5

Chapters 3 and 4 demonstrate that low frequency theta oscillations could be involved in integrating semantic content from distributed semantic networks. Yet, the discussion on

oscillatory dynamics in the context of semantic integration has started with the idea that oscillatory dynamics at different frequencies reflect network dynamics at different cortical scales. Specifically, it has been argued repeatedly that low frequency oscillations (<30 Hz) are involved in more widespread network dynamics, while high frequency oscillations (>30 Hz) are involved in local network dynamics (von Stein and Sarnthein, 2000; Donner and Siegel, 2011). In the context of semantic integration, the framework predicts that integrating more distributed cross-modal information is reflected in low frequency oscillatory dynamics, which is demonstrated in Chapter 3 and 4. However, integrating modality-specific information locally should be reflected in enhanced high frequency power.

To address this latter prediction, the experiment reported in Chapter 5 uses MEG during dual property verification. The main advantages of using MEG over EEG are that MEG usually allows for a more focal source solution than EEG. The main reason for this is that MEG is much less sensitive to secondary currents, or volume conduction (e.g., in the skin and skull), than EEG. In addition, MEG systems usually use a higher number of sensors, further improving the accuracy of the source solution. A particular advantage of MEG in the context of studying high frequency oscillations is the empirical observation that high frequency dynamics are more common in MEG than EEG studies (Muthukumaraswamy & Singh). Possible reasons for this could be that MEG is usually recorded in a magnetically shielded room, which should reduce the effects of high frequency noise sources (Baillet, 2001), and the fact that MEG is less susceptible than EEG to muscle activity (Zimmermann & Scharein, 2004).

There were three main results. First, the MEG study replicates the findings from Chapter 3 and 4, suggesting that theta oscillations reflect the integration of multimodal semantic content. Similar to the results from Chapter 3, the effect peaks in left ATL. Second, integrating features from the same modality elicits an early power increase in high frequency gamma power, which is localized to the left ATL and mPFC. Lastly, there is partial evidence for modality-specificity in auditory regions (left STG), which show a selective increase in gamma power when participants integrate two auditory features. These results confirm that integrating semantic features at different cortical scales is reflected in oscillatory neuronal dynamics at different frequencies in line with current frameworks on cortical oscillations (von Stein and Sarnthein, 2000; Donner and Siegel, 2011)

Chapter 6

Chapter 3 to 5 provide converging evidence that low frequency theta oscillations could be involved in integrating semantic content from distributed cortical networks.

Specifically, Chapter 3 demonstrates that theta oscillations connect distributed modality-specific networks to a central semantic hub in left ATL. While this is evidence that theta oscillations could be involved in linking distributed semantic information, it is not conclusive evidence that these transient connections are indeed driven by semantic content. Such a conclusion would require that the connection weights between ATL and modality-specific networks are modulated as a function of word content (Hauk & Tschentscher, 2013).

Chapter 6 investigates these predictions using an MEG dataset in which participants were asked to think about visual, auditory, or both auditory and visual features of a target word. The theta phase-locking value (PLV) between left ATL and modality-specific visual (left iOC), and auditory (left STG) networks was used as a proxy for long-range functional connectivity between these regions. The results confirm that connection weights are higher between left ATL and auditory regions when participants think about auditory versus visual features, while connections to visual areas are stronger when participants think about visual versus auditory features. Furthermore, there is a trend in the data suggesting that both connections are enhanced when participants think about visual and auditory features at the same time. These results demonstrate that theta oscillations reflect functional network interactions, which could explain how word meaning in distributed semantic networks is coordinated and eventually integrated into a coherent conceptual representation.

Selecting perceptual semantic features in memory

Research investigating the neural basis of semantic memory and word meaning, has largely focused on the question *where* information is represented in the brain. For example, functional imaging studies have demonstrated that word retrieval engages distributed cortical networks in a way that is consistent with the (perceptual) semantic features encoded in the word (Hauk et al., 2004; Kiefer et al., 2008; Martin & Chao, 2001; Simmons et al., 2007). More recent studies have shown that selecting specific features of a word can modulate the level of activation in modality-specific networks (Hoenig et al., 2008; Raposo et al., 2009; Willems and Casasanto, 2011; van Ackeren et al., 2012; van Dam et al., 2012). However, the neurophysiological processes of *how* these features are selected are still poorly understood. The present thesis addresses this problem in terms of when perceptual features are selected, and how local- and distributed neural assemblies communicate during feature selection.

Timing of lexical access

Recent evidence suggests that some semantic features are accessed even within the first 200ms after a word is presented (Hauk, Coutout, Holden, & Chen, 2012; Kiefer et al., 2008; Pulvermüller et al., 2009; Pulvermüller, Shtyrov, et al., 2005; Shtyrov, Butorina, Nikolaeva, & Stroganova, 2014). The experiments described in Chapters 2 and 5 contribute to this growing body of research. Specifically, Chapter 2 demonstrates that top-down selection of modality-specific feature content emerges already around ~220ms after a visual word is presented. Furthermore, the results from chapter 5 suggest that, around the same time period (150-350ms), the system selects multiple perceptual features in parallel, as long as all features come from the same perceptual modality (e.g., visual: green, round, shiny, etc.). These results are also in line with behavioral evidence from Chapter 3, and studies on conceptual modality switching (Pecher et al., 2003), which demonstrated processing benefits for accessing information from the same modality.

Taken together, previous work has demonstrated that dominant semantic features (e.g., motor, auditory) become available early during lexical access (Pulvermüller et al., 2005b; Kiefer et al., 2008; Shtyrov et al., 2014). Chapters 2 and 5 further show that early feature selection (around ~200ms) can be biased by the task (i.e., focus on visual, auditory, haptic features), which challenges the argument that early semantic effects reflect automatic and bottom-up processes. Furthermore, Chapter 5 reveals that around the same time window, multiple features are selected near simultaneously, as long as the information is similar (i.e., from the same modality).

Network dynamics during perceptual feature access

It has been argued that oscillatory neuronal dynamics could be used as a window to study how local- and distributed neural assemblies communicate during a cognitive task (Donner & Siegel, 2011; Hanslmayr et al., 2012; Jensen & Mazaheri, 2010; Singer & Gray, 1995; Tallon-Baudry & Bertrand, 1999; von Stein & Sarnthein, 2000). However, while oscillatory neuronal dynamics have been studied in the context of lexical-semantic retrieval before (Bastiaansen, van der Linden, Ter Keurs, Dijkstra, & Hagoort, 2005; Bastiaansen, Oostenveld, Jensen, & Hagoort, 2008; Klimesch et al., 1994; Pulvermüller, Lutzenberger, & Preissl, 1999), the rationale for studying oscillatory dynamics is often poorly motivated. This issue is discussed in some depth in Chapter 1.

Reading a word induces a complex change in oscillatory brain rhythms across multiple frequency bands. For example, as the word is presented, theta power (4-8 Hz) is enhanced, while alpha, and beta power (10-30 Hz) are suppressed (Figure 7.1). While all

experiments in the current study demonstrate a spectral profile of this form, different frequency components have been shown to be involved in different neurophysiological processes underlying word retrieval.

Chapter 2 demonstrates that beta power is suppressed when more complex information is retrieved from memory. However, no changes in beta power are observed in any of the experiments investigating feature integration (Chapter 3 and 5). Thus, while it may be that oscillatory suppression or desynchronization is sensitive to the amount of information retrieved from memory, as demonstrated previously (Khader and Rösler, 2011; Hanslmayr et al., 2012), there is no evidence that these oscillations are involved in integrating information (Bastiaansen, Magyari, & Hagoort, 2010). This might not be surprising given that the temporal correlation hypothesis states that information is integrated via synchronization rather than desynchronization (Singer and Gray, 1995). In contrast, a number of recent studies have argued that alpha/beta oscillations reflect a sustained gating of neuronal activity, which is suppressed as information becomes available (Engel & Fries, 2010; Hanslmayr et al., 2012; Jensen & Mazaheri, 2010; Jokisch & Jensen, 2007; Waldhauser et al., 2012). For example Obleser and Weisz (2012) have demonstrated, using spectrally degraded speech stimuli, that oscillatory power in the alpha/beta band is suppressed as the speech stream becomes intelligible, that is when participants start retrieving semantic content.

Taken together, the empirical work presented in the current thesis suggests that oscillatory power in the beta range might reflect an ongoing inhibition of semantic information in long-term memory, which is reduced gradually as more complex information is retrieved. It is worth pointing out that these mechanisms are by no means specific to the language system. For example, it is well established that movement initiation relies on the suppression of inhibitory mechanisms in the basal ganglia. Interestingly, local field potentials recorded from this region in Parkinson's disease reveal largely enhanced beta power that has been linked to an over-inhibition in this area (Brown, 2006).

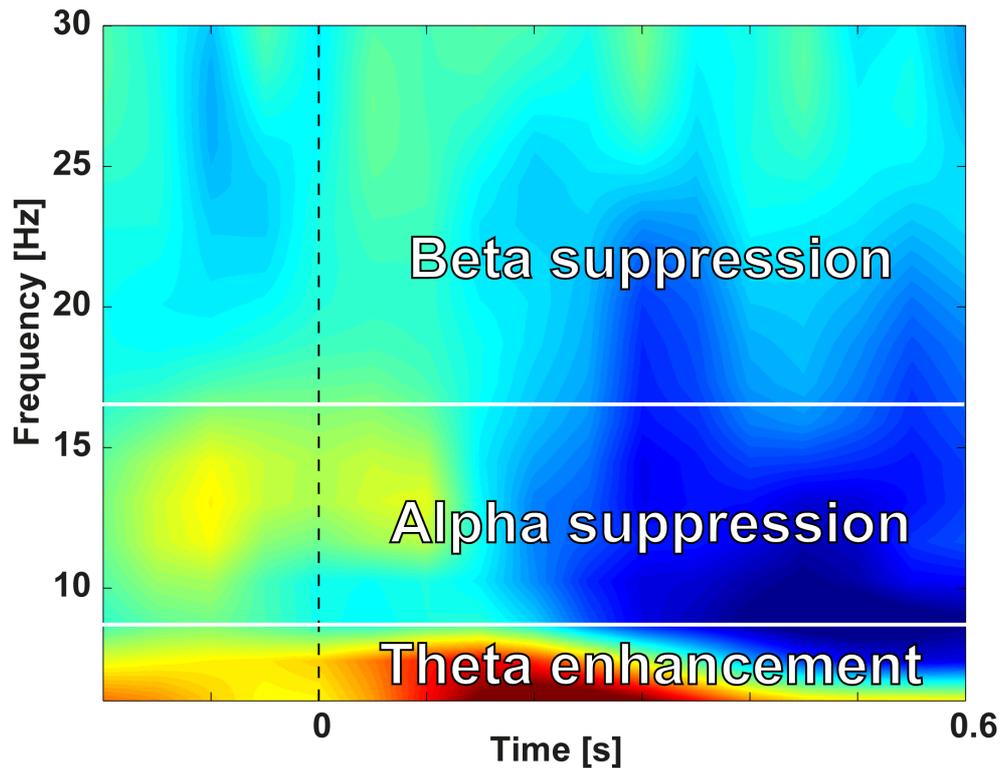


Figure 7.1. Illustration of the oscillatory dynamics during the presentation of a visual word. Low frequency-theta power shows a sustained increase, while alpha and beta power are suppressed.

Integrating multimodal semantic content

The main focus of present thesis was to investigate how multimodal semantic content is combined in a distributed semantic network. This problem can be further subdivided by asking where semantic information converges, and what the mechanism is that binds semantic content at multiple cortical scales. The next sections will discuss how the empirical work presented in the current thesis has addressed these two questions and what the implications are for neurocognitive models of semantic cognition.

Multimodal semantic content converges in left ATL

Many theories agree that semantic knowledge is represented in a distributed cortical network (Barsalou et al., 2003; Damasio, Tranel, Grabowski, Adolphs, & Damasio, 2004; Martin & Chao, 2001; Patterson et al., 2007; Vigliocco et al., 2009), however, there is little consensus about where in the brain this information is combined into more complex multimodal object representations. Based on functional imaging studies, and evidence from patients who show category selective semantic deficits as a result of stroke, some theories suggest that modality-specific semantic content is combined at

multiple locations, or convergence zones that reside in association cortices between sensory areas (Barsalou et al., 2003; Damasio, Tranel, Grabowski, Adolphs, & Damasio, 2004; Martin & Chao, 2001). In contrast, others have argued that distributed semantic content converges in a single hub, possibly in the ATL. This view is mainly based on neuropsychological evidence from dementia patients who show a generalized semantic deficit resulting from severe and progressive atrophy in the ATL (Hodges et al., 1992, 1994). However, very few imaging studies have reported ATL activation during semantic tasks, as the signal-to-noise ratio in this area is particularly low (Visser et al., 2010). Which of the two views is correct is still a major discussion in the field.

Chapter 3, 5, and 6 speak to this issue; left ATL, extending from the temporal pole ventrally into the parahippocampal gyrus, is found to be more sensitive to feature integration (Chapter 3 and 5), and shows enhanced functional connections to a distributed cortical network (Chapter 3 and 6). While other regions do show some sensitivity to this condition, only the pattern in left ATL is replicated in a different stimulus set and imaging modality (EEG in Chapter 3 and 4, and MEG in Chapter 5). Recent studies have demonstrated that conceptual processing in the ATL is partly reliant on perceptual processing in posterior areas (Coutanche and Thompson-Schill, 2014), and targeting left ATL with theta burst stimulation disrupts binding between an object and its canonical color (Chiou et al., 2013). However, the majority of these studies are restricted to visual image processing. The current study uniquely contributes to this body of evidence demonstrating that left ATL fulfills a similar function when semantic content of words is retrieved from long-term memory.

But, while much of the evidence in Chapter 3 and 5 suggests that left ATL is involved in semantic feature integration, Chapter 5 demonstrates that integrating features from the same modality also engages modality-specific semantic networks. For example, integrating two auditory features (e.g., *loud*, *shrill*) with a target word activates an auditory network in left STG. Taken together, these results suggest that semantic features are not integrated at one location only (ATL). Rather, depending on the perceptual semantic content, graded feature integration could already occur in modality-specific networks.

Lastly, it should be pointed out that, although the results from Chapter 3, 5 and 6 support the idea of a central hub in left ATL, this does not mean that convergence zones are not relevant for semantic cognition. Indeed, the current thesis suggests that integrating features from the same modality already starts in modality-specific networks. The hub might be an additional layer where information from multiple modalities converges. There is currently a divide between theories that advocate a single hub, versus theories in favor of multiple convergence zones. Yet, the current thesis

demonstrates that the answer might lie in the combination of both approaches. However, one difficulty in comparing the present results to previous work is that many previous studies have used images rather than words as stimuli, and there is evidence to suggest that category specific patterns of activation might be sensitive to the presentation modality (Devlin, 2002; Devereux et al., 2013; Rice et al., 2014).

Oscillatory dynamics reflect local- and long-range network dynamics during semantic retrieval.

The question where in the brain semantic information is combined has received a lot of attention in recent years. However, few studies have investigated what the network dynamics within a central hub could be, and how it communicates with distributed modality-specific networks. The experiments described in Chapters 3 to 6 directly address this problem by investigating changes in oscillatory activity when participants combine semantic feature from the same or multiple modalities. Oscillatory dynamics are thought to reflect synchronized firing of large neuronal populations (Singer and Gray, 1995; Musall et al., 2012), and are considered a proxy for local (von Stein and Sarnthein, 2000; Donner and Siegel, 2011), and long-range network interactions (Hipp, Engel, & Siegel, 2011; Klimesch, Freunberger, Sauseng, & Gruber, 2008; Lachaux et al., 1999; Varela & Lachaux, 2001).

Cortical oscillations are interesting for studying how semantic information is dynamically integrated at multiple levels. For example, local oscillatory changes (changes in spectral power) should respond differently to information that is more distributed at the level of the cortex (e.g. information from the same, or different modalities). This is an advantage with respect to functional imaging in which the dependent variable is referred to as activation or percent signal change (Singh, 2012). In addition, phase relationships between oscillatory dynamics in different regions are often used as a proxy for how multiple regions communicate (Lachaux et al., 1999). In the context of semantic processing, this method can be used to test how retrieving different types of information affects the connection weights between a single hub, or multiple convergence zones and modality-specific cortical networks.

The experiments described in Chapters 3 to 5 consistently demonstrate that low frequency theta oscillations, generated in left ATL, are involved in integrating perceptual semantic features; all three experiments show a more sustained increase in the theta band when participants are asked to integrate features from multiple modalities, that is, features that are represented in a more distributed fashion. Importantly, these results should not be interpreted such that theta oscillations are sensitive to cross-modal features pairs only; rather, theta oscillations could be generally involved in dynamically

linking neural assemblies in the temporal lobe to distributed sites in the neocortex (Hipp et al., 2012). Specifically, Chapter 3 demonstrates that theta oscillations in left ATL and a distributed cortical network become phase-locked when participants retrieve, and integrate feature knowledge from the same or multiple modalities. The central role of theta oscillations for temporal lobe connectivity has been demonstrated in a recent study in resting state networks (Hipp et al., 2012). Hipp and colleagues demonstrated that the connectivity pattern in the temporal lobe reveals hub-like network properties (high degrees, high betweenness) particularly in the theta range. However, extending these findings, and the results from Chapter 3, Chapter 6 demonstrates that theta oscillations connecting left ATL to modality-specific networks are modulated by semantic feature content. Thus, the empirical work presented in previous chapters suggests that theta oscillations reflect the coordination of distributed information from long-term memory, which is an inherent component of semantic integration. Furthermore, these dynamics likely reflected both in theta power, and phase.

While theta oscillations seem to be involved in long-range network interactions, Chapter 5 demonstrates that gamma oscillations could be particularly sensitive to local network interactions. Specifically, integrating features from the same modality has been shown to elicit enhanced gamma activity in left ATL as well as modality-specific cortical networks (left STG for auditory feature integration). These results suggest a complex system of high and low frequency oscillations that retrieve, and possibly integrate, information at different cortical scales (von Stein and Sarnthein, 2000; Donner and Siegel, 2011). As discussed in the introduction, enhanced theta and gamma power are frequently observed in memory research (Axmacher et al., 2006; Osipova et al., 2006; Düzel et al., 2010). However, the studies presented in the current thesis are unique in the way that the contribution of each frequency is linked to the scale at which semantic information is integrated.

Nevertheless, a final question remains: how do local and long-range network dynamics, or high, and low oscillatory frequencies interface with each other? Previous research has demonstrated that local gamma power is modulated, or gated by the phase of the theta, and in some experiments alpha, rhythm during memory tasks (Axmacher et al., 2010; Canolty and Knight, 2010; Voytek et al., 2010; Fries et al., 2012). The analyses in Chapter 5 and 6 provide indirect evidence for this claim. Specifically areas, which showed a local increase in gamma power during modality-specific feature integration (left STG, and iOC), also revealed enhanced phase-locking with left ATL in a way that was consistent with the semantic content of the features.

Towards a neurophysiological model of semantic feature retrieval and integration

The experiments presented in the current thesis provide unique, and novel insights into the neurophysiological basis of semantic feature selection and integration. These findings should be considered a starting point for future research, and a guide to formulate hypotheses about language, and semantic processing, which are in line with current frameworks on cortical oscillations. The model described in Figure 7.2 is an attempt to integrate the current findings with the hub-and-spokes model, which proved to be most consistently supported by the experiments in the present thesis (Chapter 3 and 5).

The main architecture of the model consists of modality specific semantic networks that are connected via a central hidden layer in the ATL. The current findings suggest that these long-range connections are reflected in slow theta oscillations (Chapter 3 and 6). In contrast local interactions in modality-specific regions are characterized by high-frequency gamma, and possibly beta oscillations (Chapter 2 and 5). However, while gamma oscillations are shown to be sensitive to the integration of information (Chapter 5), beta suppression might reflect the read-out of semantic features from memory (Chapter 2). However, as no source reconstruction was performed on the data in Chapter 2, whether beta suppression is modality-specific needs to be further investigated. There is some prior research suggesting that this might be the case (Jokisch and Jensen, 2007; Khader and Rösler, 2011; Waldhauser et al., 2012). Lastly, Chapter 5 reports that gamma oscillations in the ATL are also sensitive to local feature integration. Modality-specificity in the ATL has only recently been demonstrated (Bonner & Price, 2013; Peelen & Caramazza, 2012), and certainly merits further investigation. The finding that different oscillatory dynamics are sensitive to different types of network interactions, even within the same region, can be a powerful way to isolate specific sub-processes in the ATL.

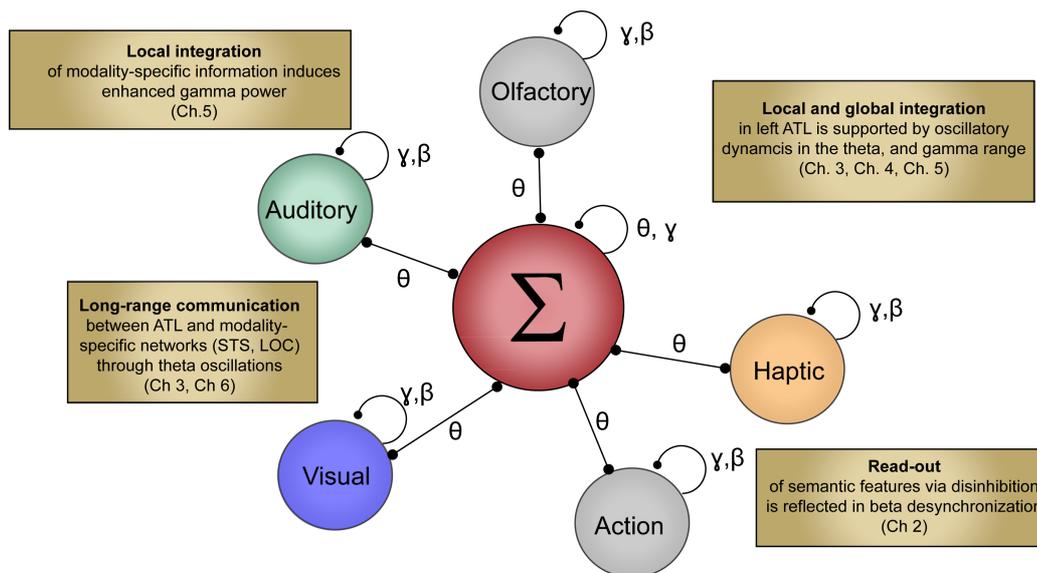


Figure 7.2. The extended hub-and-spokes model. The foundation of the model is the hub-and-spokes model in which modality-specific semantic networks (green, blue, orange, grey) are connected via a central hub in the ATL (red). The extended model includes parameters illustrating which oscillatory frequencies operate at different levels of the model. Long-range connections are reflected in low frequency theta oscillations, while local interactions (illustrated as recurrent connections) are dominated by high-frequency gamma oscillations. Beta suppression could reflect the readout of modality-specific feature content.

Language as a model for testing theories on cortical oscillations

The major aim of the current thesis was to investigate the neurophysiological basis of semantic feature selection and integration. To this end, oscillatory neuronal dynamics were used as a tool to study local- and long-range network interactions during semantic tasks. However, the experimental work presented in the current thesis reveals that it is possible to use language as a tool to test hypotheses about the function of different cortical oscillations. For example, Chapter 2 demonstrates that the complexity of the information retrieved is related to oscillatory desynchronization which is directly in line with the information via de-synchronization hypothesis (Hanslmayr et al., 2012). Furthermore, the dissociation between high, and low frequency oscillations (Chapter 3 and 5) for integrating local and distributed information is support for a recent framework by Donner and Siegel (2011), and an earlier hypothesis by von Stein and Sarnthein (2000). In short, despite earlier inconsistencies in studies investigating oscillatory dynamics during language processing, the experiments in the current thesis demonstrate

that it is possible to address questions about the function of different oscillatory dynamics using language processing and vice versa.

Limitations

It is argued throughout the thesis that the study of cortical oscillations can provide novel and unique insights into the neurophysiological basis of semantic processing, and the underlying network dynamics in particular. Yet, there are a number of more general limitations that require further clarification.

The time scale of oscillatory responses

Recent evidence suggests that modality-specific semantic knowledge is accessed very rapidly (Hauk et al., 2012; Kiefer et al., 2008; Shtyrov et al., 2014), and this is demonstrated in Chapter 2 of the current thesis. Yet, some of the effects in the frequency domain (e.g., theta power increase for cross-modal feature integration) emerge as late as 750ms after target onset. There are a number of reasons why the timing of oscillatory changes is difficult to interpret. First, the time-frequency analyses performed in the current thesis, particularly in the low frequency domain, are computed using a sliding time window of 500ms. While this is necessary for a robust estimation of spectral power in the low frequency domain, as well as a sufficient frequency resolution, it also introduces severe time-domain smoothing. Thus a late oscillatory component is likely to start much earlier in time.

Second, one might argue that low frequency oscillatory dynamics are too slow to capture fast language phenomena. It is demonstrated in Chapter 2 that low frequency cortical oscillations in the first 500ms connect left ATL to a widespread cortical network across all conditions. Given the speed and fluency of language processing, it may be that the network is triggered already in early stages of visual processing, as a stimulus is interpreted or predicted to be a word. From this point onwards, even if it took a full theta cycle, modality-specific content could still be accessed between 125 and 250ms after stimulus onset (for a 4-8 theta rhythm). However, while activating the network is likely to be fast and locked to stimulus onset, the offset is much more variable. Indeed it has been demonstrated that cortical rhythms, induced by alternating current stimulation (TACS), show an after effect beyond stimulation offset (Neuling et al., 2013). In the context of the experiments in Chapters 3-5, the difference between integrating cross modal versus modality-specific features observed around 750ms likely emerges at an earlier point in time. Again, the timing of the effect is very difficult to interpret, and should not directly be compared with data from evoked potentials.

Lastly, it should be emphasized that the goal of studying oscillatory dynamics should not be to answer questions about the timing of an effect, but rather about the neurophysiological mechanisms, or network dynamics

The preparation period and pre-target integration

In the dual property verification tasks described in Chapters 3-6, it could be argued that some form of integration already occurs before the target word is presented. However, while participants will certainly have made predictions about the upcoming stimulus during that period, the stimuli were designed, and rated, such that the target word was unlikely to be guessed based on the feature pair alone. In addition none of the experiments showed a significant effect in the period preceding the target word. Thus, the moment of integration was most likely reduced to the presentation of the target word.

Recommendations for future research

The current thesis has demonstrated that oscillatory dynamics can be used to acquire novel insights into the neurophysiological mechanisms underlying semantic feature selection and integration. However, as pointed out previously in this chapter, the framework presented here should not be considered exhaustive. Rather, it can be used as a guide for future studies on language processing. In this section, I will point out some of the ways in which the current framework could be used or refined.

In recent years, a large number of functional imaging studies have demonstrated how MVPA can be used to test whether a region is sensitive to semantic differences between conditions (e.g., Peelen and Caramazza, 2012; Coutanche and Thompson-Schill, 2014). Other studies, particularly in the embodied cognition literature, have used conjunction analysis to show that the same region is sensitive to different tasks (e.g., color perception and color words) (Simmons et al., 2007; Kiefer et al., 2008). However, these studies cannot be considered evidence that two tasks also engage the same neuronal populations. In contrast, chapter 5 demonstrates that the same region might be sensitive to different conditions, but the oscillatory dynamics are still very different. Thus, oscillatory dynamics might be a useful way of insulating different processes even within the same region. This could be particularly important for a focused investigation of what is encoded in the ATL.

Furthermore, fMRI studies are usually interpreted in terms of task induced activation in a given region. Yet, Chapter 2 has demonstrated that semantic features

might be selected through disinhibition, rather than activation. It is not clear how these patterns relate to the neural activity usually reported in fMRI studies (Singh, 2012).

The experiments presented in Chapter 3 to 6 suggest that integrating semantic features is supported by an interplay between high and low frequency oscillations. It has been argued in recent years that theta reflects a gating mechanism such that gamma power is more enhanced at specific points in the theta cycle. However, while phase-amplitude coupling has been demonstrated in memory tasks, much of this evidence is based on electrocorticography (ECoG) data in patients who are undergoing surgery for epilepsy (Voytek et al., 2010). Due to a much lower signal to noise ratio, as well as the lack of a source reconstruction method that truthfully captures this effect, evidence for phase-amplitude coupling in EEG and MEG is rare. Thus, future studies using ECoG recordings could use the present framework to test if the theta rhythm in ATL does in fact modulate gamma amplitude in modality-specific regions.

Lastly, all experiments presented in the current thesis demonstrate oscillatory changes that co-occur with a particular task manipulation. Similar to fMRI, these data should be considered correlational. In other words, there is no evidence that cortical oscillations drive semantic processing in a causal way. However, recent studies have started to artificially induce oscillatory frequencies that resemble the natural cortical rhythms of the brain (Neuling et al., 2013). This method, called transcranial alternating current stimulation (TACS), has great potential for testing whether oscillatory frequencies have a causal role in cognitive processes. For example, based on the current data one might predict that inducing a theta rhythm in the ATL could activate the semantic network and improve performance on semantic tasks. In contrast, inducing a beta rhythm should block features selection, and impair semantic processing significantly.

Conclusions

A large number of studies in recent years, using a range of different methods, have suggested that word meaning might at least be partially stored in distributed modality-specific networks (Barsalou et al., 2003; Binder & Desai, 2011; Damasio et al., 2004; Martin & Chao, 2001). Furthermore, the level of activation in each of these networks depends on the dominant semantic features of a word (e.g., action verbs activate action areas) (Hauk et al., 2004; Kiefer et al., 2008; Simmons et al., 2007), and what type of information the person is currently attending to (e.g., do you use the object with your hand?) (Hoenig et al., 2008; van Dam et al., 2012). The aim of the present thesis was to investigate a) how these features are selected and b) how multiple features are combined

into a coherent representation. The main focus of the current work was on the role of cortical oscillations in these processes. Specifically, cortical oscillations in recent years have been used as a proxy for different types of network dynamics in the underlying neuronal populations (Donner and Siegel, 2011). The current results suggest that top-down selection of semantic features starts early (around ~200ms), and likely involves a selective disinhibition, which is reflected in a reduction in oscillatory power around the beta band. In contrast, feature integration involves an interplay between local- and long-range network dynamics that are reflected in low frequency theta, and high frequency gamma power. Finally, the results suggest that the left ATL might play a critical role in binding semantic information, by linking semantic content from distributed cortical networks.

Appendix

Complete lists of all stimuli used in the experiments described in Chapter 2-6 can be retrieved from the website accompanying this thesis:

<https://sites.google.com/site/oscillationslexicalsemantics/>

Bibliography

- Aristotle, Hamlyn DW, Shields CJ (1993) *De anima* : books II and III with passages from book I. Clarendon Press ;Oxford University Press.
- Axmacher N, Henseler MM, Jensen O, Weinreich I, Elger CE, Fell J (2010) Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proc Natl Acad Sci U S A* 107:3228–3233.
- Axmacher N, Mormann F, Fernández G, Elger CE, Fell J (2006) Memory formation by neuronal synchronization. *Brain Res Rev* 52:170–182.
- Baillet S, Moshier JC, Leahy RM (2001) Electromagnetic brain mapping. *Networks* 18:14–30.
- Barsalou LW (1999) Perceptual symbol systems. *Behav Brain Sci* 22:577–609; discussion 610–60.
- Barsalou LW (2008) Grounded cognition. *Annu Rev Psychol* 59:617–645.
- Barsalou LW, Pecher D, Zeelenberg R, Simmons WK, Hamann SB (2005) Multimodal simulation in conceptual processing. In: *Categorization inside and outside the lab: Festschrift in honor of Douglas L. Medin* (Ahn W, Goldstone R, Love B, Markman A, Wolff P, eds), pp 249–270. Washington, DC: American Psychological Association.
- Barsalou LW, Simmons K, Barbey AK, Wilson CD (2003) Grounding conceptual knowledge in modality-specific systems. *Trends Cogn Sci* 7:84–91.
- Bastiaansen M, Magyari L, Hagoort P (2010) Syntactic unification operations are reflected in oscillatory dynamics during on-line sentence comprehension. *J Cogn Neurosci* 22:1333–1347.
- Bastiaansen MC, van der Linden M, Ter Keurs M, Dijkstra T, Hagoort P (2005) Theta responses are involved in lexical-semantic retrieval during language processing. *J Cogn Neurosci* 17:530–541.

- Bastiaansen MCM, 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 Lang* 106:15–28.
- Bastiaansen MCM, van Berkum JJ a, Hagoort P (2002) Event-related theta power increases in the human EEG during online sentence processing. *Neurosci Lett* 323:13–16.
- Beauchamp MS, Martin A (2007) Grounding object concepts in perception and action: evidence from fMRI studies of tools. *Cortex* 43:461–468.
- Berger H (1929) Über das elektrenkephalogramm des menschen. *Eur Arch Psychiatry Clin*
- Binder JR, Desai RH (2011) The neurobiology of semantic memory. *Trends Cogn Sci* 15:527–536.
- Binder JR, Desai RH, Graves WW, Conant LL (2009) Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex* 19:2767–2796.
- Bonner MF, Grossman M (2012) Gray matter density of auditory association cortex relates to knowledge of sound concepts in primary progressive aphasia. *J Neurosci* 32:7986–7991.
- Bonner MF, Price AR (2013) Where is the anterior temporal lobe and what does it do? *J Neurosci* 33:4213–4215.
- Boulenger V, Hauk O, Pulvermüller F (2009) Grasping ideas with the motor system: semantic somatotopy in idiom comprehension. *Cereb Cortex* 19:1905–1914.
- Boulenger V, Mechtouff L, Thobois S, Broussolle E, Jeannerod M, Nazir TA (2008) Word processing in Parkinson's disease is impaired for action verbs but not for concrete nouns. *Neuropsychologia* 46:743–756.
- Brewer JB, Zhao Z, Desmond JE, Glover GH, Gabrieli JD (1998) Making memories: brain activity that predicts how well visual experience will be remembered. *Science* 281:1185–1187.
- Brown P (2006) Bad oscillations in Parkinson's disease. *J Neural Transm Suppl*:27–30.

- Brysbaert M, New B, Keuleers E (2012) Adding part-of-speech information to the SUBTLEX-US word frequencies. *Behav Res Methods* 44:991–997.
- Buccino G, Riggio L, Melli G, Binkofski F, Gallese V, Rizzolatti G (2005) Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. Elsevier.
- Burgess AP, Gruzelier JH (2000) Short duration power changes in the EEG during recognition memory for words and faces.
- Buzsáki G (2005) Theta rhythm of navigation: link between path integration and landmark navigation, episodic and semantic memory. *Hippocampus* 15:827–840.
- Canolty RT, Knight RT (2010) The functional role of cross-frequency coupling. *Trends Cogn Sci* 14:506–515.
- Catani M, Thiebaut de Schotten M (2008) A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex* 44:1105–1132.
- Cervenka MC, Nagle S, Boatman-Reich D (2011) Cortical High-Gamma Responses in Auditory Processing. *Am J Audiol* 20:171–181.
- Chatterjee A (2010) Disembodying cognition. *Lang Cogn* 2.
- Chiou R, Sowman P, Etchell A, Rich A (2013) A Conceptual Lemon: Theta Burst Stimulation to the Left Anterior Temporal Lobe Untangles Object Representation and Its Canonical Color. :1066–1074.
- Connell L, Lynott D (2011) Modality Switching Costs Emerge in Concept Creation as Well as Retrieval. *Cogn Sci* 35:763–778.
- Coutanche MN, Thompson-Schill SL (2014) Creating Concepts from Converging Features in Human Cortex. *Cereb Cortex*.
- Creem SH, Proffitt DR (2001) Grasping objects by their handles: a necessary interaction between cognition and action. *J Exp Psychol Hum Percept Perform* 27:218–228.
- Damasio AR (1989) The brain binds entities and events by multiregional activation from convergence zones. *Neural Comput* 1:123–132.

- Damasio H, Grabowski T, Tranel D (1996a) A neural basis for lexical retrieval. *Nature*.
- Damasio H, Grabowski TJ, Tranel D, Hichwa RD, Damasio AR (1996b) A neural basis for lexical retrieval. *Nature* 380:499–505.
- Damasio H, Tranel D, Grabowski T, Adolphs R, Damasio AR (2004) Neural systems behind word and concept retrieval. *Cognition* 92:179–229.
- Davidson DJ, Indefrey P (2007) An inverse relation between event-related and time-frequency violation responses in sentence processing. *Brain Res* 1158:81–92.
- Devereux BJ, Clarke A, Marouchos A, Tyler LK (2013) Representational similarity analysis reveals commonalities and differences in the semantic processing of words and objects. *J Neurosci* 33:18906–18916.
- Devlin JT (2002) Is there an anatomical basis for category-specificity? Semantic memory studies in PET and fMRI. *Neuropsychologia* 40:54–75.
- Dils AT, Boroditsky L (2010) Visual motion aftereffect from understanding motion language. *Proc Natl Acad Sci U S A* 107:16396–16400.
- Donner TH, Siegel M (2011) A framework for local cortical oscillation patterns. *Trends Cogn Sci* 15:191–199.
- Donoghue JP, Sanes JN, Hatsopoulos NG, Gaál G (1998) Neural discharge and local field potential oscillations in primate motor cortex during voluntary movements. *J Neurophysiol* 79:159–173.
- Düzel E, Penny WD, Burgess N (2010) Brain oscillations and memory. *Curr Opin Neurobiol* 20:143–149.
- Engel AK, Fries P (2010) Beta-band oscillations-signalling the status quo? *Curr Opin Neurobiol* 20:156–165.
- Engel AKK, Fries P, Singer W, others (2001) Dynamic predictions: oscillations and synchrony in top-down processing. *Nat Rev Neurosci* 2:704–716.
- Fell J, Axmacher N (2011) The role of phase synchronization in memory processes. *Nat Rev Neurosci* 12:105–118.

- Fodor J (1981) The modularity of mind.
- Fox PT, Mikiten S, Davis G, Lancaster JL (1994) BrainMap: A database of human function brain mapping. In: *Functional Neuroimaging: Technical Foundations* (Thatcher R.W., Hallett M, Zeffiro T, John ER, Huerta M, eds), pp 95–105. San Diego, California: Academic Press.
- Foxe JJ, Simpson G V, Ahlfors SP (1998) Parieto-occipital approximately 10 Hz activity reflects anticipatory state of visual attention mechanisms.
- Foxe JJ, Snyder AC (2011) The Role of Alpha-Band Brain Oscillations as a Sensory Suppression Mechanism during Selective Attention. *Front Psychol* 2:154.
- Fries P (2009) Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annu Rev Neurosci* 32:209–224.
- Friese U, Köster M, Hassler U, Martens U, Trujillo-Barreto N, Gruber T (2012) Successful memory encoding is associated with increased cross-frequency coupling between frontal theta and posterior gamma oscillations in human scalp-recorded EEG. *Neuroimage* 66C:642–647.
- Gauthier I, Anderson AW, Tarr MJ, Skudlarski P, Gore JC (1997) Levels of categorization in visual recognition studied using functional magnetic resonance imaging. *Curr Biol* 7:645–651.
- Gentilucci M, Gangitano M (1998) Influence of automatic word reading on motor control. *Neuroscience* 10:752–756.
- Glenberg AM, Kaschak MP (2002) Grounding language in action. *Psychon Bull Rev* 9:558–565.
- Goldberg RF, Perfetti CA, Schneider W (2006a) Perceptual knowledge retrieval activates sensory brain regions. *J Neurosci* 26:4917–4921.
- Goldberg RF, Perfetti CA, Schneider W (2006b) Perceptual knowledge retrieval activates sensory brain regions. *J Neurosci* 26:4917–4921.
- González J, Barros-Loscertales A, Pulvermüller F, Meseguer V, Sanjuán A, Belloch V, Avila C (2006) Reading cinnamon activates olfactory brain regions. *Neuroimage* 32:906–912.

- Gross J, Baillet S, Barnes GR, Henson RN, Hillebrand A, Jensen O, Jerbi K, Litvak V, Maess B, Oostenveld R, Parkkonen L, Taylor JR, van Wassenhove V, Wibral M, Schoffelen J-M (2012) Good practice for conducting and reporting MEG research. *Neuroimage*.
- Gross J, Kujala J, Hamalainen M, Timmermann L, Schnitzler A, Salmelin R, Hämäläinen M (2001) Dynamic Imaging of Coherent Sources: Studying Neural Interactions in the Human Brain. *Proc Natl Acad Sci* 98:694–699.
- Guderian S, Düzel E (2005) Induced theta oscillations mediate large-scale synchrony with mediotemporal areas during recollection in humans. *Hippocampus* 15:901–912.
- Hagoort P (2005) On Broca, brain, and binding: a new framework. *Trends Cogn Sci* 9:416–423.
- Hagoort P, Hald L, Bastiaansen M, Petersson KM (2004) Integration of Word Meaning and World Knowledge in Language Comprehension. *Science* (80-) 304:438–441.
- Hald L a, Bastiaansen MCM, Hagoort P (2006) EEG theta and gamma responses to semantic violations in online sentence processing. *Brain Lang* 96:90–105.
- Hanslmayr S, Staudigl T, Fellner M-C (2012) Oscillatory power decreases and long-term memory: the information via desynchronization hypothesis. *Front Hum Neurosci* 6:74.
- Hari R, Forss N, Avikainen S, Kirveskari E, Salenius S, Rizzolatti G (1998) Activation of human primary motor cortex during action observation: a neuromagnetic study. *Proc Natl Acad Sci U S A* 95:15061–15065.
- Harnad S (1990) The symbol grounding problem. 42.
- Hauk O, Coutout C, Holden A, Chen Y (2012) The time-course of single-word reading: Evidence from fast behavioral and brain responses. *Neuroimage* 60:1462–1477.
- Hauk O, Johnsrude I, Pulvermüller F, Pulvermuller F (2004) Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41:301–307.
- Hauk O, Tschentscher N (2013) The Body of Evidence: What Can Neuroscience Tell Us about Embodied Semantics? *Front Psychol* 4:50.

- Hipp JF, Engel AK, Siegel M (2011) Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron* 69:387–396.
- Hipp JF, Hawellek DJ, Corbetta M, Siegel M, Engel AK (2012) Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nat Neurosci* 15:884–890.
- Hodges JR, Patterson K, Oxbury S, Funnell E (1992) Semantic dementia: progressive fluent aphasia with temporal lobe atrophy. *Brain* 115:1783–1806.
- Hodges JR, Patterson K, Tyler LK (1994) Loss of semantic memory: implications for the modularity of mind. *Cogn Neuropsychol* 11:505–542.
- Hoenig K, Sim E-J, Bochev V, Herrnberger B, Kiefer M (2008) Conceptual flexibility in the human brain: dynamic recruitment of semantic maps from visual, motor, and motion-related areas. *J Cogn Neurosci* 20:1799–1814.
- Jefferies E, Lambon Ralph MA (2006) Semantic impairment in stroke aphasia versus semantic dementia: a case-series comparison. *Brain* 129:2132–2147.
- Jensen O, Gelfand J, Kounios J, Lisman JE (2002) Oscillations in the alpha band (9-12 Hz) increase with memory load during retention in a short-term memory task. *Cereb Cortex* 12:877–882.
- Jensen O, Kaiser J, Lachaux JP (2007) Human gamma-frequency oscillations associated with attention and memory. *Trends Neurosci* 30:317–324.
- Jensen O, Mazaheri A (2010) Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Front Hum Neurosci* 4:186.
- Jensen O, Tesche CD (2002) Frontal theta activity in humans increases with memory load in a working memory task. *Eur J Neurosci* 15:1395–1399.
- Jokisch D, Jensen O (2007) Modulation of gamma and alpha activity during a working memory task engaging the dorsal or ventral stream. *J Neurosci* 27:3244–3251.
- Kaiser J, Lutzenberger W, Ackermann H, Birbaumer N (2002) Dynamics of gamma-band activity induced by auditory pattern changes in humans. *Cereb Cortex* 12:212–221.

- Khader PH, Rösler F (2011) EEG power changes reflect distinct mechanisms during long-term memory retrieval. *Psychophysiology* 48:362–369.
- Kiefer M, Sim E-J, Herrnberger B, Grothe J, Hoenig K (2008) The sound of concepts: four markers for a link between auditory and conceptual brain systems. *J Neurosci* 28:12224–12230.
- Klimesch W (1999) EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res Brain Res Rev* 29:169–195.
- Klimesch W, Doppelmayr M, Yonelinas a, Kroll NE, Lazzara M, Röhlm D, Gruber W (2001) Theta synchronization during episodic retrieval: neural correlates of conscious awareness. *Brain Res Cogn Brain Res* 12:33–38.
- Klimesch W, Freunberger R, Sauseng P (2010) Oscillatory mechanisms of process binding in memory. *Neurosci Biobehav ...* 34:1002–1014.
- Klimesch W, Freunberger R, Sauseng P, Gruber W (2008) A short review of slow phase synchronization and memory: Evidence for control processes in different memory systems? *Brain Res* 1235:31–44.
- Klimesch W, Schimke H, Schwaiger J (1994) Episodic and semantic memory: an analysis in the EEG theta and alpha band. *Electroencephalogr Clin Neurophysiol* 91:428–441.
- Kounios J, Green DL, Payne L, Fleck JI, Grondin R, McRae K (2009) Semantic richness and the activation of concepts in semantic memory: Evidence from event-related potentials. *Brain Res* 1282:95–102.
- Kutas M, Federmeier K (2000) Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn Sci* 4:463–470.
- Kutas M, Hillyard SA (1980) Reading senseless sentences: brain potentials reflect semantic incongruity. *Science (80-)* 207:203–205.
- Lachaux JP, Rodriguez E, Martinerie J, Varela FJ (1999) Measuring phase synchrony in brain signals. *Hum Brain Mapp* 8:194–208.

- Lambon Ralph M a, Sage K, Jones RW, Mayberry EJ (2010) Coherent concepts are computed in the anterior temporal lobes. *Proc Natl Acad Sci U S A* 107:2717–2722.
- Landauer TK, Foltz PW, Laham D (1998) An introduction to latent semantic analysis. *Discourse Process* 25:259–284.
- Liljeström M, Kujala J, Jensen O, Salmelin R (2005) Neuromagnetic localization of rhythmic activity in the human brain: a comparison of three methods. *Neuroimage* 25:734–745.
- Lisman J (2010) Working memory: the importance of theta and gamma oscillations. *Curr Biol* 20:R490–2.
- Lynott D, Connell L (2009) Modality exclusivity norms for 423 object properties. *Behav Res Methods* 41:558–564.
- Mahon BZ, Caramazza A (2008) A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J Physiol Paris* 102:59–70.
- Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164:177–190.
- Marr D (1982) *Vision*. book.
- Martin A, Chao LL (2001) Semantic memory and the brain: structure and processes. *Curr Opin Neurobiol* 11:194–201.
- McClelland JL, Rogers TT, others (2003) The parallel distributed processing approach to semantic cognition. *Nat Rev Neurosci* 4:310–322.
- McNorgan C, Reid J, McRae K (2011) Integrating conceptual knowledge within and across representational modalities. *Cognition* 118:211–233.
- Meteyard L, Cuadrado SR (2010) Coming of age: A review of embodiment and the neuroscience of semantics. ... 48:788–804.
- Meteyard L, Vigliocco G (2008) The role of sensory and motor information in semantic representation: A review. *Handb Cogn Sci An embodied approach San Diego Elsevier*:293–312.

- Milner PM (1967) A model for visual shape recognition. *Psychol Rev* 81:521–535 ST –
A model for visual shape recognition.
- Moreno I, De Vega M, Leon I (2013) Understanding action language modulates
oscillatory mu and beta rhythms in the same way as observing actions. *Brain Cogn*
82:236–242.
- Musall S, von Pföstel V, Rauch A, Logothetis NK, Whittingstall K (2012) Effects of
Neural Synchrony on Surface EEG. *Cereb Cortex*:1–9.
- Muthukumaraswamy SD (2010) Functional Properties of Human Primary Motor Cortex
Gamma Oscillations. *J Neurophysiol* 104:2873–2885.
- Muthukumaraswamy SD, Singh KD (2013) Visual gamma oscillations: The effects of
stimulus type, visual field coverage and stimulus motion on MEG and EEG
recordings. *Neuroimage* 69:223–230.
- Neininger B, Pulvermüller F (2003) Word-category specific deficits after lesions in the
right hemisphere. *Neuropsychologia* 41:53–70.
- Neuling T, Rach S, Herrmann CS (2013) Orchestrating neuronal networks: sustained
after-effects of transcranial alternating current stimulation depend upon brain
states. *Front Hum Neurosci* 7:161.
- Neuper C, Wörtz M, Pfurtscheller G (2006) ERD/ERS patterns reflecting sensorimotor
activation and deactivation. *Prog Brain Res* 159:211–222.
- Niedermeyer E, Lopes da Silva F (1993) *Electroencephalography: Basic Principles,
Clinical Applications, and Related Fields*. Williams & Wilkins, Baltimore.
- Nolte G (2003) The magnetic lead field theorem in the quasi-static approximation and
its use for magnetoencephalography forward calculation in realistic volume
conductors. *Phys Med Biol* 48:3637–3652.
- Nolte G, Bai O, Wheaton L, Mari Z, Vorbach S, Hallett M (2004) Identifying true brain
interaction from EEG data using the imaginary part of coherency. *Clin
Neurophysiol* 115:2292–2307.
- O’Keefe J, Recce ML (1993) Phase relationship between hippocampal place units and
the EEG theta rhythm. *Hippocampus* 3:317–330.

- Obleser J, Weisz N (2012) Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cereb Cortex* 22:2466–2477.
- Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:1.
- Oostenveld R, Stegeman DF, Praamstra P, Van Oosterom A (2003) Brain symmetry and topographic analysis of lateralized event-related potentials. *Clin Neurophysiol* 114:1194–1202.
- Osipova D, Takashima A, Oostenveld R, Fernández G, Maris E, Jensen O (2006) Theta and Gamma Oscillations Predict Encoding and Retrieval of Declarative Memory. *J Neurosci* 26:7523–7531.
- Panagiotaropoulos TI, Deco G, Kapoor V, Logothetis NK (2012) Neuronal Discharges and Gamma Oscillations Explicitly Reflect Visual Consciousness in the Lateral Prefrontal Cortex. *Neuron* 74:924–935.
- Papeo L, Vallesi A, Isaja A, Rumiati RI (2009) Effects of TMS on Different Stages of Motor and Non-Motor Verb Processing in the Primary Motor Cortex Ferrari PF, ed. *PLoS One* 4:11.
- Patterson K, Nestor PJ, Rogers TT (2007) Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat Rev Neurosci* 8:976–987.
- Pecher D, Zeelenberg R, Barsalou LW (2002) Verifying Properties from Different Modalities for Concepts Produces Switching Costs.
- Pecher D, Zeelenberg R, Barsalou LW (2003) Verifying different-modality properties for concepts produces switching costs.
- Peelen M V, Caramazza A (2012) Conceptual object representations in human anterior temporal cortex. *J Neurosci* 32:15728–15736.
- Pexman PM, Holyk GG, Monfils M-H (2003) Number-of-features effects and semantic processing. *Mem Cognit* 31:842–855.

- Pexman PM, Lupker SJ, Hino Y (2002) The impact of feedback semantics in visual word recognition: number-of-features effects in lexical decision and naming tasks. *Psychon Bull Rev* 9:542–549.
- Pfurtscheller G, Lopes da Silva FH (1999) Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin Neurophysiol* 110:1842–1857.
- Pfurtscheller G, Neuper C (2010) Dynamics of Sensorimotor Oscillations in a Motor Task. In: Graimann B, Pfurtscheller G, Allison B, eds. *Potentials*:47–64.
- Plaut DC (2002a) Graded modality-specific specialization in semantics: a computational account of optic aphasia. *Cogn Neuropsychol* 19:603–639.
- Plaut DC (2002b) Graded modality-specific specialisation in semantics: A computational account of optic aphasia.
- Pobric G, Jefferies E, Ralph M a L (2010a) Amodal semantic representations depend on both anterior temporal lobes: evidence from repetitive transcranial magnetic stimulation. *Neuropsychologia* 48:1336–1342.
- Pobric G, Jefferies E, Ralph MAL (2010b) Report Category-Specific versus Category-General Semantic Impairment Induced by Transcranial Magnetic Stimulation. *Curr Biol* 20:964–968.
- Postle N, McMahon KL, Ashton R, Meredith M, de Zubicaray GI (2008) Action word meaning representations in cytoarchitecturally defined primary and premotor cortices. *Neuroimage* 43:634–644.
- Pulvermüller F (2013) How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. *Trends Cogn Sci* 17:458–470.
- Pulvermüller F, Fadiga L (2010) Active perception: sensorimotor circuits as a cortical basis for language. *Nat Rev Neurosci* 11:351–360.
- Pulvermüller F, Hauk O (2006) Category-specific conceptual processing of color and form in left fronto-temporal cortex. *Cereb Cortex* 16:1193–1201.
- Pulvermüller F, Hauk O, Nikulin V V, Ilmoniemi RJ (2005a) Functional links between motor and language systems. *Eur J Neurosci* 21:793–797.

- Pulvermüller F, Lutzenberger W, Preissl H (1999) Nouns and Verbs in the Intact Brain: Evidence from Event-Related Potentials and High-Frequency Cortical Responses. *Cereb Cortex* 9:497–506.
- Pulvermüller F, Shtyrov Y (2009) Spatiotemporal signatures of large-scale synfire chains for speech processing as revealed by MEG. *Cereb Cortex* 19:79–88.
- Pulvermüller F, Shtyrov Y, Hauk O (2009) Understanding in an instant: neurophysiological evidence for mechanistic language circuits in the brain. *Brain Lang* 110:81–94.
- Pulvermüller F, Shtyrov Y, Ilmoniemi R (2005b) Brain signatures of meaning access in action word recognition.
- Rabovsky M, Sommer W, Abdel Rahman R (2012) The time course of semantic richness effects in visual word recognition. *Front Hum Neurosci* 6.
- Raghavachari S, Kahana MJ, Rizzuto DS, Caplan JB, Kirschen MP, Bourgeois B, Madsen JR, Lisman JE (2001) Gating of human theta oscillations by a working memory task. *J Neurosci* 21:3175–3183.
- Raghavachari S, Lisman JE, Tully M, Madsen JR, Bromfield EB, Kahana MJ (2006) Theta oscillations in human cortex during a working-memory task: evidence for local generators. *J Neurophysiol* 95:1630–1638.
- Raposo A, Moss HE, Stamatakis EA, Tyler LK (2009) Modulation of motor and premotor cortices by actions, action words and action sentences. *Neuropsychologia* 47:388–396.
- Rice GE, Watson DM, Hartley T, Andrews TJ (2014) Low-Level Image Properties of Visual Objects Predict Patterns of Neural Response across Category-Selective Regions of the Ventral Visual Pathway. *J Neurosci* 34:8837–8844.
- Richter T, Zwaan R a (2009) Processing of color words activates color representations. *Cognition* 111:383–389.
- Rogers TT, Lambon Ralph M a, Garrard P, Bozeat S, McClelland JL, Hodges JR, Patterson K (2004) Structure and deterioration of semantic memory: a neuropsychological and computational investigation. *Psychol Rev* 111:205–235.

- Rosenblum LD (2008) Speech Perception as a Multimodal Phenomenon. *Curr Dir Psychol Sci* 17:405–409.
- Rueschemeyer S, Lindemann O, van Elk M, Bekkering H (2009) Embodied cognition: The interplay between automatic resonance and selection-for-action mechanisms. *Eur J Soc Psychol* 39:1180–1187.
- Rueschemeyer S-A, Glenberg AM, Kaschak MP, Mueller K, Friederici AD (2010a) Top-down and bottom-up contributions to understanding sentences describing objects in motion. *Front Psychol* 1:183.
- Rueschemeyer S-A, van Rooij D, Lindemann O, Willems RM, Bekkering H (2010b) The function of words: distinct neural correlates for words denoting differently manipulable objects. *J Cogn Neurosci* 22:1844–1851.
- Rüschemeyer S-A, Brass M, Friederici AD (2007) Comprehending prehending: neural correlates of processing verbs with motor stems. *J Cogn Neurosci* 19:855–865.
- Sauseng P, Griesmayr B, Freunberger R, Klimesch W (2010) Control mechanisms in working memory: a possible function of EEG theta oscillations. *Neurosci Biobehav Rev* 34:1015–1022.
- Schneider TR, Debener S, Oostenveld R, Engel AK (2008a) Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming. *Neuroimage* 42:1244–1254.
- Schneider TR, Engel AK, Debener S (2008b) Multisensory Identification of Natural Objects in a Two-Way Crossmodal Priming Paradigm. *Exp Psychol (formerly Zeitschrift für Exp Psychol* 55:121–132.
- Schneider TR, Lorenz S, Senkowski D, Engel AK (2011) Gamma-band activity as a signature for cross-modal priming of auditory object recognition by active haptic exploration. *J Neurosci* 31:2502–2510.
- Senkowski D, Schneider TR, Foxe JJ, Engel AK (2008) Crossmodal binding through neural coherence: implications for multisensory processing. *Trends Neurosci* 31:401–409.

- Shannon CE, Weaver W (1949) *The Mathematical Theory of Communication* (Shannon CE, Weaver W, eds). University of Illinois Press.
- Shtyrov Y, Butorina A, Nikolaeva A, Stroganova T (2014) Automatic ultrarapid activation and inhibition of cortical motor systems in spoken word comprehension. *Proc Natl Acad Sci U S A*:1918–1923.
- Simmons WK, Ramjee V, Beauchamp MS, McRae K, Martin A, Barsalou LW (2007) A common neural substrate for perceiving and knowing about color. *Neuropsychologia* 45:2802–2810.
- Singer W, Gray CM (1995) Visual Feature Integration and the Temporal Correlation Hypothesis. *Annu Rev Neurosci* 18:555–586.
- Singh KD (2012) Which “neural activity” do you mean? fMRI, MEG, oscillations and neurotransmitters. *Neuroimage* 62:1121–1130.
- Sober SJ, Sabes PN (2003) Multisensory integration during motor planning. *J Neurosci* 23:6982–6992.
- Spitzer B, Hanslmayr S, Opitz B, Mecklinger A, Bäuml K-H (2009) Oscillatory correlates of retrieval-induced forgetting in recognition memory. *J Cogn Neurosci* 21:976–990.
- Summerfield C, Mangels J a (2005) Coherent theta-band EEG activity predicts item-context binding during encoding. *Neuroimage* 24:692–703.
- Tallon-Baudry C, Bertrand O (1999) Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn Sci* 3:151–162.
- Tallon-Baudry C, Bertrand O, Delpuech C, Permier J (1997a) Oscillatory gamma-band (30-70 Hz) activity induced by a visual search task in humans. *J Neurosci* 17:722–734.
- Tallon-Baudry C, Bertrand O, Wienbruch C, Ross B, Pantev C (1997b) Combined EEG and MEG recordings of visual 40 Hz responses to illusory triangles in human.
- Tettamanti M, Buccino G, Saccuman MC, Gallese V, Danna M, Scifo P, Fazio F, Rizzolatti G, Cappa SF, Perani D (2005) Listening to action-related sentences activates fronto-parietal motor circuits. *J Cogn Neurosci* 17:273–281.

- Tomasino B, Fink GR, Sparing R, Dafotakis M, Weiss PH (2008) Action verbs and the primary motor cortex: a comparative TMS study of silent reading, frequency judgments, and motor imagery. *Neuropsychologia* 46:1915–1926.
- Tranel D, Damasio H, Damasio AR (1997) A neural basis for the retrieval of conceptual knowledge. *Neuropsychologia* 35:1319–1327.
- Treisman A (1996) The binding problem. *Curr Opin Neurobiol* 6:171–178.
- Van Ackeren MJ, Casasanto D, Bekkering H, Hagoort P, Rueschemeyer S-A (2012) Pragmatics in Action: Indirect Requests Engage Theory of Mind Areas and the Cortical Motor Network. *J Cogn Neurosci* 24:2237–2247.
- Van Ackeren MJ, Rueschemeyer S-A (2014) Cross-modal integration of lexical-semantic features during word processing: Evidence from oscillatory dynamics during EEG. *PLoS One*.
- Van Ackeren MJ, Schneider TR, Müsch K, Rueschemeyer S-A (in press) Oscillatory neuronal activity reflects lexical-semantic feature integration within and across sensory modalities in distributed cortical networks. *The Journal of Neuroscience*.
- Van Dam WO, Rueschemeyer S-A, Bekkering H (2010) How specifically are action verbs represented in the neural motor system: an fMRI study. *Neuroimage* 53:1318–1325.
- Van Dam WO, van Dijk M, Bekkering H, Rueschemeyer S-A (2012) Flexibility in embodied lexical-semantic representations. *Hum Brain Mapp* 33:2322–2333.
- Van Dantzig S, Cowell RA, Zeelenberg R, Pecher D (2010) A sharp image or a sharp knife: norms for the modality-exclusivity of 774 concept-property items. *Behav Res Methods* 43:145–154.
- Van Der Werf J, Jensen O, Fries P, Medendorp WP (2008) Gamma-band activity in human posterior parietal cortex encodes the motor goal during delayed prosaccades and antisaccades. *J Neurosci* 28:8397–8405.
- Van Elk M, Van Schie HT, Zwaan RA, Bekkering H (2010) The functional role of motor activation in language processing: motor cortical oscillations support lexical-semantic retrieval. *Neuroimage* 50:665–677.

- Varela F, Lachaux J (2001) The brainweb: phase synchronization and large-scale integration. *Nat Rev ...* 2:229–239.
- Vigliocco G, Meteyard L, Andrews M, Kousta S (2009) Toward a theory of semantic representation. *Lang Cogn* 1:219–247.
- Visser M, Jefferies E, Lambon Ralph M a (2010) Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. *J Cogn Neurosci* 22:1083–1094.
- Von der Malsburg C, Schneider W (1986) A neural cocktail-party processor. *Biol Cybern* 54:29–40.
- Von Stein A, Sarnthein J (2000) Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. *Int J Psychophysiol* 38:301–313.
- Voytek B, Canolty RT, Shestyuk A, Crone NE, Parvizi J, Knight RT (2010) Shifts in gamma phase-amplitude coupling frequency from theta to alpha over posterior cortex during visual tasks. *Front Hum Neurosci* 4:191.
- Waldhauser GT, Johansson M, Hanslmayr S (2012) Alpha/Beta Oscillations Indicate Inhibition of Interfering Visual Memories. *J Neurosci* 32:1953–1961.
- Wang L, Jensen O, van den Brink D, Weder N, Schoffelen J-M, Magyari L, Hagoort P, Bastiaansen M (2012a) Beta oscillations relate to the N400m during language comprehension. *Hum Brain Mapp* 00.
- Wang L, Zhu Z, Bastiaansen M (2012b) Integration or Predictability? A Further Specification of the Functional Role of Gamma Oscillations in Language Comprehension. *Front Psychol* 3:187.
- Warrington EK (1975) The selective impairment of semantic memory. *Q J Exp Psychol* 27:635–657.
- Willems RM, Casasanto D (2011) Flexibility in embodied language understanding. *Front Psychol* 2:116.

- Womelsdorf T, Schoffelen J-M, Oostenveld R, Singer W, Desimone R, Engel AK, Fries P (2007) Modulation of Neuronal Interactions Through Neuronal Synchronization. *Science* (80-) 316:1609–1612.
- Wu X, Chen X, Li Z, Han S, Zhang D (2007) Binding of verbal and spatial information in human working memory involves large-scale neural synchronization at theta frequency. *Neuroimage* 35:1654–1662.
- Zimmermann R, Scharein E (2004) MEG and EEG show different sensitivity to myogenic artifacts. *Neurol Clin Neurophysiol* 2004:78.
- Zwaan R a, Stanfield R a, Yaxley RH (2002) Language comprehenders mentally represent the shapes of objects. *Psychol Sci* 13:168–171.
- Zwaan RA (2003) The Immersed Experiencer: Toward An Embodied Theory Of Language Comprehension. *Psychol Learn Motiv - Adv Res Theory* 44:35–62.