Semantic Richness Effects in Visual Word Processing: Combining Event-Related Brain Potentials and Connectionist Network Modeling

Dissertation

zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)
im Fach Psychologie

eingereicht an der Mathematisch-Naturwissenschaftlichen Fakultät II
der Humboldt-Universität zu Berlin

von Dipl.-Psych. Milena Rabovsky

Dekan: Prof. Dr. Elmar Kulke

Gutachter/innen: 1. Prof. Dr. Rasha Abdel Rahman

2. Prof. Dr. Werner Sommer

3. Prof. Dr. Markus Kiefer

Datum der Verteidigung: 15. Juli 2013

Table of contents

Acknowledgements

First of all, I wish to thank Rasha Abdel Rahman and Werner Sommer for their highly supportive supervision. My work greatly benefitted from their positive and encouraging attitude, their advice and curiosity, their extraordinary reliability and speed, their constant support in developing and following my interests, and last but not least their friendliness and humanity - thank you very much! I am also very grateful to Ken McRae for his kind hospitality and support, and for giving me the opportunity to work on a modeling project in his lab despite being a complete novice.

Many thanks to Markus Kiefer for reviewing this thesis. It is a great pleasure having my work reviewed by an expert whose research contributed to first sparking my interest in cognitive neuroscience.

I enjoyed the cooperative and friendly atmosphere in the Biological Psychology and Neuro-cognitive Psychology labs, where I could always be sure to get answers and help when needed, and therefore wish to thank both teams. I thank the Berlin School of Mind and Brain for funding, and for providing an inspiring environment for completing this thesis. I am particularly grateful to Felix Blankenburg for his friendliness and expertise, especially for pointing me to predictive coding which substantially shaped my thinking about the results of this thesis. Many thanks to my fellow students for sharing important and not so important things in life and science, and for the good time.

I am very grateful to my friends for their support and patience over the years. Thank you for being around. I thank Daniel for being so encouraging and helpful, for his interest in my work that resulted in many extremely enjoyable and insightful discussions and much advice, for his emotional support, and for being there. Finally, I wish to thank my parents for their support in many respects, and for their confidence.

Abstract

Language ultimately aims to convey meaning. Importantly, the amount of associated semantic information varies considerably between individual words. Recent evidence suggests that the richness of semantic representations can facilitate performance in lexical and semantic tasks, but much remains to be learned about semantic richness effects. The present dissertation combined event-related brain potentials (ERPs) and connectionist network modeling to address several unresolved issues concerning the role of semantic richness in word processing. Specifically, ERPs were employed to investigate the time course of independent influences of the number of semantic features and associates during word reading (study 1) and influences of semantic richness on implicit word learning (study 2). Aiming at advancing a mechanistic understanding of the obtained results, both studies were subsequently simulated using a network model of semantic cognition (study 3).

Results showed no influences of the number of associates, but fast access to semantic features, with influences of feature-based semantic richness starting at about 190 ms - a mere 20 to 30 ms after and temporally overlapping with the activation of orthographic representations as reflected by N1 lexicality effects. Later on, a higher number of semantic features induced larger N400 amplitudes. Furthermore, the number of semantic features enhanced repetition priming effects on lexical decision accuracy and N400 amplitudes, providing initial evidence for influences of semantic richness on implicit word learning. These results are in line with feature-based network model of semantic cognition. Simulations with such a model suggested that semantic activation can facilitate lexical decisions, while network error closely corresponds to N400 amplitudes. In psychological terms, network error has been conceptualized as implicit prediction error. Thus, these results are taken to suggest that N400 amplitudes may reflect implicit prediction error in semantic memory.

Zusammenfassung

Lesen zielt darauf ab, Bedeutung aus geschriebenem Text zu extrahieren. Interessanterweise unterscheiden sich Wörter beträchtlich hinsichtlich der Menge mit ihnen assoziierter Bedeutung, und es wurde kürzlich gezeigt, dass eine hohe Bedeutungshaltigkeit lexikalische und semantische Aufgaben erleichtern kann. Die vorliegende Dissertation kombiniert ereigniskorrelierte Hirnpotentiale (EKPs) und konnektionistische Netzwerk-Modellierung, um einige offene Fragen zur Rolle der Bedeutungshaltigkeit bei der Wortverarbeitung anzugehen. Hierbei wurden EKPs verwendet, um den Zeitverlauf unabhängiger Einflüsse der Anzahl semantischer Merkmale und Assoziationen beim Wortlesen zu bestimmen sowie Einflüsse von Bedeutungshaltigkeit auf implizites Wortlernen zu untersuchen. Um die zugrundeliegenden Mechanismen besser zu verstehen, wurden die Ergebnisse anschließend mittels eines semantischen Netzwerk-Modells simuliert.

Es zeigten sich keine Einflüsse der Anzahl der Assoziationen, aber eine schnelle Aktivierung semantischer Merkmale, die das EKP bereits ab 190 ms beeinflussten - nur 20 bis 30 ms nach und zeitlich überlappend mit der Aktivierung orthographischer Repräsentationen, die durch N1-Lexikalitätseffekte angezeigt wurden. Im weiteren Verlauf ging eine hohe Merkmalsanzahl mit größeren N400-Amplituden einher. Zudem verstärkten semantische Merkmale Wiederholungseinflüsse auf die Akkuratheit lexikalischer Entscheidungen und N400-Amplituden, was einen ersten Hinweis auf Einflüsse von Bedeutungshaltigkeit auf implizites Wortlernen darstellt. Diese Ergebnisse stehen im Einklang mit merkmalsbasierten semantischen Netzwerk-Modellen. Simulationen legen nahe, dass semantische Aktivierung lexikalische Entscheidungen erleichtert, während Netzwerk-Fehler in engem Zusammenhang mit N400-Amplituden stehen. Da Netzwerk-Fehler psychologisch als implizite Vorhersagefehler interpretiert werden, deuten diese Ergebnisse darauf hin, dass N400-Amplituden implizite Vorhersagefehler im semantischen System widerspiegeln könnten.

Synopsis

1 Introduction

Getting along in daily life depends heavily on vast amounts of knowledge about the objects around us – what we might find inside the white and rectangular object in the kitchen, and what we could do with the things we encounter; how the four-legged furry animal on the street might react when we tried to touch it, how that touch would feel, etc. Such knowledge is often referred to as semantic knowledge and is assumed to be represented in semantic memory. Representations in semantic memory endow sensory input with meaning, thereby enabling us to interact with the things surrounding us in a reasonable way and to make sense of spoken or written words.

Interestingly, words differ widely in the amount of associated semantic information. For instance, the semantic representation associated with the word ‘car’ is multifaceted and rich (cars have doors, windows, wheels, engines; require gasoline; cause pollution; are used for transportation, as a status symbol, etc.), while the semantic representation linked to the word ‘cork’ is relatively sparse (corks are cylindric, small, used to close wine bottles). As language ultimately aims to convey meaning, the richness of semantic representations might be expected to have an important impact on word processing. Indeed, recent evidence suggests that semantic richness can facilitate performance in lexical and semantic tasks (see below for further discussion). However, much remains to be learned about semantic richness effects and their contribution to understanding the nature of semantic representations and the mechanisms underlying the activation of these representations during reading.

With the aim of contributing to elucidate these issues, the present dissertation linked time-resolved measurements of electrical brain activity as provided by event-related brain potentials (ERPs) and connectionist modeling, two approaches which have both brought about substantial progress in understanding semantic cognition (see below), but have very rarely been combined. Specifically, ERPs were employed to investigate the temporal evolvement of semantic richness effects during word reading (study 1) and influences of semantic richness on implicit word learning (study 2). Aiming at advancing a mechanistic understanding of the obtained results, specifically the modulations of the meaning-related N400 ERP component, both studies were subsequently simulated using a connectionist network model of semantic cognition (study 3). The next sections provide some relevant background on the investigation of semantic cognition, focusing on semantic network models (section 1.1), semantic richness effects (section 1.2), and meaning-related ERPs (section 1.3). The three studies are subsequently summarized in sections 2, and jointly discussed in section 3.

1.1 Models of semantic cognition

Researchers have long been interested in the representation and computation of meaning. Early views conceived of semantic memory as an amodal, modular store of factual information (Tulving, 1972); a predominant theory on how knowledge could be represented in this system assumed a large set of stored propositions (Collins & Quillian, 1969). Further developments within the general frame of an amodal, modular semantic system included the proposal of feature-based representations of category prototypes (Rosch & Mervis, 1975), which arose based on evidence for the gradedness of category membership (e.g. sperlings are much more typical birds than penguins), and the so-called theory-theory which held that the subtlety and sophistication of semantic judgments could only be explained by implicit causal theories structuring semantic knowledge (Gopnik & Wellmann, 1994).

Mechanistic understanding of semantic processes has been considerably advanced by the explicit computational implementation of models of semantic cognition. Especially network models from the connectionist approach have been extremely successful in explaining a wide range of phenomena in semantic cognition (including those that had motivated complex theoretical assumptions such as the theory-theory, hierarchical representations, or stored prototypes) as arising from the domain-general learning principles of the connectionist framework (Cree, McNorgan, & McRae, 2006; Cree, McRae, & McNorgan, 1999; O'Connor, Cree, & McRae, 2009; Rogers & McClelland, 2008). In this frame, semantic cognition is suggested to arise from activation flowing among simple processing units according to the strength of the weights connecting the units. Semantic knowledge is assumed to be stored in these connection weights which are gradually adjusted by experience. In response to a perceptual input, such as a word or an object, the semantic system is assumed to make information available which is not directly present in the input. Thus, upon seeing a dog or a bird, a person with a fully developed semantic system can roughly anticipate how it would sound if the respective animal would make a sound or what would happen if a cat would come around the corner. Similarly, when reading sentences such as “I take my coffee with cream and …” the semantic system generates anticipations of plausible continuations. Semantic representations are assumed to be shaped by adjusting the connection weights supporting these anticipations based on prediction errors, i.e. the difference between the anticipations and actual outcomes, resulting in more accurate anticipations in the future. In this way, such implicit predictions (which may be weak and random in newborn babies) can gradually improve through experience.

In connectionist models, such learning corresponds to the activation of a specific input pattern (representing perceptual input) upon which activation flowing according to connection weights produces a specific output pattern (representing implicit predictions) which is compared to a target pattern (representing actual outcomes). Connection weights are then adjusted based on the difference between model-generated output and target (representing implicit predictions and actual outcomes, respectively), by means of error-backpropagation.

Such implicit error-driven learning is assumed to occur incessantly and automatically during information processing, in order to gain an accurate internal representation of the environment and optimize processing. It should be noted that the assumption that prediction errors drive learning is not specific to the connectionist framework, but widely shared in cognitive science and neuroscience (e.g. McLaren, 1989; Schultz & Dickinson, 2000; Tobler, O'Doherty, Dolan, & Schultz, 2006), and has recently been compellingly advanced by predictive coding accounts of the neuro-cognitive system (den Ouden, Daunizeau, Roiser, Friston, & Stephan, 2010; den Ouden, Friston, Daw, McIntosh, & Stephan, 2009; Friston, 2005, 2009). The possibility to understand semantic cognition in terms of such general mechanisms, with knowledge stored in connections shaped by experiences, challenges assumptions of a strict separation and modularity of semantic memory, instead suggesting a more interactive and integrated view of the semantic system as dynamically interwoven in experience. Such a perspective seems nicely in line with the proposal that semantic cognition is grounded in perception and action, with semantic features stored in perceptual and motor systems according to sensory-motor experiences during concept acquisition (Barsalou, 1999; Kiefer & Spitzer, 2001; Pulvermüller, 2005).

A promising development within the connectionist modeling approach is to derive the semantic information to be learned by the models empirically, instead of based on intuition (Andrews, Vigliocco, & Vinson, 2009; Cree, et al., 2006; O'Connor, et al., 2009). Such an approach allows for a better approximation of human semantic representations, thereby improving model predictions, and additionally reduces degrees of freedom for the modeler. Thus, a number of models (including the one used in study 3) have implemented semantic representations based on the elaborate feature production norms by McRae et al. (2005) where more than 700 participants listed 2526 semantic features (e.g. mouse – “is small”, “has four legs” etc.) for 541 concrete words (Cree, et al., 2006; O'Connor, et al., 2009). Another recent model implemented semantic representations based on a combination of these features with word co-occurrences in text corpora (Andrews, et al., 2009).

Even though it seems beneficial to use empirical information (e.g. feature norms or co-occurrence statistics) as a basis for implementing semantic representations, it is of course a crucial issue what kind of empirical information the implementation should be based on. Although many theories and models assume semantic features to play an important role in the representation of meaning (Collins & Loftus, 1975; Harm & Seidenberg, 2004; McRae, deSa, & Seidenberg, 1997; Plaut & Shallice, 1993; Rogers & McClelland, 2008), it is an ongoing debate whether the organization of semantic memory is based on semantic features or associations or both (Hutchison, 2003; Lucas, 2000; Yee, Overton, & Thompson-Schill, 2009). Measures of the richness of semantic representations may help to elucidate this issue, as discussed in the next section.

1.2 Semantic richness

As noted above, words differ considerably in the amount of associated semantic information (e.g. ‘car’ vs. ‘cork’). Extensive norming efforts during recent years have provided a well-grounded empirical basis to systematically quantify this variance along various dimensions, and to explore the influences of these various measures, such as for example the number of features generated in the above-described feature production norms (McRae, et al., 2005), the number of associates, referring to the number of different first associations generated across participants in free-association tasks (Nelson, McEvoy, & Schreiber, 2004), the diversity of contexts in which a word appears (Adelman, Brown, & Quesada, 2006), or word co-occurrences in text corpora (Buchanan, Westbury, & Burgess, 2001).

Recent evidence suggests facilitative influences of semantic richness in visual word processing (Adelman, et al., 2006; Buchanan, et al., 2001; Dunabeitia, Aviles, & Carreiras, 2008; Pexman, Holyk, & Monfils, 2003; Pexman, Lupker, & Hino, 2002) and free recall (Hargreaves, Pexman, Johnson, & Zdrazilova, 2012), and semantic richness has been shown to induce decreased activation in various brain regions (Pexman, Hargreaves, Edwards, Henry, & Goodyear, 2007). However, many issues remain unresolved. First, different dimensions of semantic richness are often positively correlated and current work just starts to disentangle independent contributions of different measures (Pexman, Hargreaves, Siakaluk, Bodner, & Pope, 2008; Yap, Pexman, Wellsby, Hargreaves, & Huff, 2012; Yap, Tan, Pexman, & Hargreaves, 2011). Such disentanglement is important because independent influences of a specific measure of semantic richness suggest that the underlying dimension indeed plays a role in semantic processing. Thus, this kind of evidence helps to elucidate the structure of semantic representations and is therefore crucial to the above-discussed issue concerning what kind of empirical information should be used to best approximate the nature of semantic representations in implemented models. In this spirit, study 1 aimed to disentangle independent influences of the number of semantic features and associates during word reading.

Second, evidence concerning the time course of semantic richness effects is scarce. Only a few previous studies have exploited the high temporal resolution provided by ERPs (see next section) to investigate semantic richness effects, and the obtained results are rather inconsistent (Amsel, 2011; Kounios et al., 2009; Müller, Dunabeitia, & Carreiras, 2010). However, such evidence is important to understand the mechanisms underlying influences of semantic richness and their interplay with other processes involved in visual word recognition. Thus, study 1 employed ERPs to further investigate the temporal dynamics of semantic richness effects.

Furthermore, while evidence concerning semantic richness effects on word recognition rapidly accumulates, little is known about whether and how the richness of semantic representations may influence language learning. Study 2 aimed at sheding some initial light on this question by investigating influences of the number of semantic features on implicit word learning as measured by repetition effects on lexical decision performance and the N400 component of the ERP, which provides a well-established electrophysiological indicator of semantic processing, as described below.

1.3 Electrophysiological indicators of semantic cognition

Event-related brain potentials (ERPs) measure the electrical activity of the brain (primarily summed post-synaptic potentials of synchronously activated neurons in the neocortex) related to an event, for example a visual word. The continuous online signal provided by ERPs seems to be very well-suited to investigate semantic processing not only due to its high temporal resolution but also because semantic computations are often not trivially related to behavioral performance (Cree, et al., 2006). Many lexical or semantic tasks depend on very specific decision thresholds to be crossed (Grondin, Lupker, & McRae, 2009), which may prevent those aspects of semantic processing that are irrelevant to these decision thresholds from being reflected in behavioral performance. ERP measures thus seem ideally suited to inform conceptions of continuous internal processes such as those implemented in the above-discussed network models of semantic cognition.

Recent studies have found semantic variables to modulate the ERP as early as within the first 200 ms of word processing (Dambacher, Rolfs, Gollner, Kliegl, & Jacobs, 2009; Hauk, Davis, Ford, Pulvermuller, & Marslen-Wilson, 2006; Hoenig, Sim, Bochev, Herrnberger, & Kiefer, 2008; Kiefer, Sim, Herrnberger, Grothe, & Hoenig, 2008; Penolazzi, Hauk, & Pulvermüller, 2007; Rabovsky, Sommer, & Abdel Rahman, 2012a; Skrandies, 1998). Nonetheless, the most well-established and widely used electrophysiological indicator of semantic cognition is the N400 component, a negative ERP component with a broad centro-parietal scalp distribution peaking at about 400 ms after the presentation of a potentially meaningful stimulus (Kutas & Federmeier, 2011; Kutas & Hillyard, 1980). The N400 has been consistently linked to semantic processing in a multitude of paradigms, with for example larger amplitudes for semantic violations (Kutas & Hillyard, 1980), low cloze probability sentence continuations (Kutas & Hillyard, 1984), and targets following a semantically unrelated as compared to a related prime, be it a word, a face, or a sound (Barrett & Rugg, 1989; Bentin, Mccarthy, & Wood, 1985; Van Petten & Rheinfelder, 1995). Important for present purposes, N400 amplitudes have also been shown to be modulated by repetition, with smaller amplitudes for repeated stimuli (Nagy & Rugg, 1989), and by the richness of semantic representations. However, the direction of the obtained semantic richness effects differed between experiments, with two studies reporting larger N400 amplitudes for words with richer semantic representations (Amsel, 2011; Müller, et al., 2010), while Kounios et al. (2009) observed a trend in the opposite direction. Clearly, further research seems desirable.

In general, N400 data have yielded important insights into semantic processing, e.g. pointing at a more interactive (Kutas, 1993) and proactive (Federmeier & Kutas, 1999) semantic system than previously assumed, and blurring psycholinguistic distinctions between meaning at different levels (e.g. word meaning vs. world knowledge; Hagoort, Hald, Bastiaansen, & Petersson, 2004). However, despite the large body of data, there is currently no agreement on the specific computational mechanisms underlying the N400. The theoretical debate on the functional basis of the N400 mostly consists in verbally descriptive proposals such as semantic binding (Federmeier & Laszlo, 2009), semantic memory access (Kutas & Federmeier, 2000; Lau, Phillips, & Poeppel, 2008), semantic integration (Baggio & Hagoort, 2011), or semantic inhibition (Debruille, 2007), which are difficult to unequivocally identify with specific measures in mechanistic theoretical frameworks such as the above-described models of semantic cognition. Unfortunately, even though the continuity of the ERP signal seems to relate very nicely to the continuous internal processes simulated in network models of semantic cognition, there has been very little contact between ERPs and computational modeling in research on semantic processing (see Laszlo & Plaut, 2012, for a very recent exception).

1.4 Aims and outline of the present work

The present work aims at narrowing this gap, combining ERPs and connectionist modeling to further elucidate semantic richness effects in visual word processing. The dissertation is composed of three studies, two ERP studies which were subsequently simulated in a modeling study. Stimuli for the ERP studies were selected from the feature production norms by McRae et al. (2005) which were also the basis for semantic representations in the network model, enabling direct item-specific simulation. The ERP studies investigated the time course of independent influences of the number of semantic features and associates during word reading (study 1; section 2.1) and influences of feature-based semantic richness on implicit word learning (study 2; section 2.2). The modeling study primarily addressed the functional basis of the observed N400 amplitude modulations, but also examined the mechanisms underlying influences of semantic richness on lexical decision performance (study 3; section 2.3).

2 Summary of the present studies

2.1 The time course of semantic richness effects during word reading (study 1)

An important step towards understanding the mechanisms underlying the extraction of meaning from print is to investigate which of the various measures proposed to quantify semantic richness influence word processing, and at what point in time these influences take place (absolutely as well as in relation to other lexical variables). In study 1, we thus contrasted two important measures of the richness of semantic representations, the number of semantic features (McRae, et al., 2005) and associates (Nelson, et al., 2004), as described in the introduction. 160 word stimuli were selected so that the number of semantic features and associates were orthogonally manipulated in the stimulus set. ERPs were recorded while these words were presented along with an equal number of pseudowords, and participants performed lexical decisions. We did not find any influence of the number of associates. In contrast, the number of semantic features modulated ERP amplitudes starting at about 190 ms already. Aiming to specify the temporal relationship between word form and meaning processing, we additionally compared the onset of semantic richness effects with the onset of lexicality effects: as pseudowords do not match any pre-existing visual word form representation, ERP differences between words and pseudowords may already arise at the level of orthographic processing, preceding possible effects at the semantic level. Lexicality effects began at about 164 ms in the left-lateralized N1 component, suggested to reflect visual word form processing (McCandliss, Cohen, & Dehaene, 2003), and continued while semantic feature effects set in. Thus, influences of semantic features started a mere 20 to 30 ms after, and temporally overlapping with, form-related processes during reading. Later on, in the N400 segment, the number of semantic features enhanced negative amplitudes at centro-parietal sites. This N400 effect is at variance with results of Kounios et al. (2009) who found a trend for larger N400 amplitudes for words with few semantic features, so that further research seems desirable. On the other hand, the present finding is in line with more negativity for words with many features in the N400 segment as reported by Amsel (2011), as well as the finding that concrete words (often assumed to be associated with richer semantic representations) induce larger N400 amplitudes than abstract words (Holcomb, Kounios, Anderson, & West, 1999; Kounios, et al., 2009; Kounios & Holcomb, 1994; West & Holcomb, 2000), and that newly learned objects and their written names elicit larger N400 amplitudes when they are associated with in-depth as compared to minimal semantic information (Abdel Rahman & Sommer, 2008; Rabovsky, et al., 2012a). The finding of feature effects arising already at about 190 ms is in line with previous evidence for fast access to semantics in reading (Dambacher, et al., 2009; Hauk, et al., 2006; Kiefer, et al., 2008; Penolazzi, et al., 2007; Rabovsky, et al., 2012a; Skrandies, 1998), and quite clearly converges with earlier suggestions that N400 effects occur too late to reflect the first phase of lexical semantic access (e.g. Dambacher, et al., 2009; Hauk, et al., 2006), stimulating the question of the functional basis of the observed N400 modulation. This issue was left unresolved in study 1 (please see Rabovsky, Sommer, & Abdel Rahman, 2012c, p. 8), but will receive further consideration in study 3 and the discussion section. Summing up, initial access to semantic features is fast, taking place within the first 200 ms, and feature-based semantic richness continues to modulate processing later on during reading. In contrast, we did not observe any influence of the number of associates.

2.2 Influences of semantic richness on implicit word learning (study 2)

While by now quite a few studies suggest that semantic richness facilitates visual word recognition (Grondin, et al., 2009; Pexman, et al., 2003; Pexman, et al., 2002), and a very recent study showed that semantic richness enhances free recall (Hargreaves, et al., 2012), as yet little is known about whether and how the richness of semantic representations may influence implicit learning and adaptation processes. Study 2 explored this issue by investigating influences of the amount of associated semantic features on implicit word learning.

Implicit learning occurs incidentally during information processing, and it is often assumed that prediction errors play an important role in this process. As described in the introduction, it has been suggested that the brain incessantly and automatically anticipates upcoming events based on an experience-derived internal model of the environment. Deviations between anticipated and factual events are assumed to drive adaptations of internal representations to reduce future prediction errors and optimize processing (Friston, 2009; McClelland, 1994; McLaren, 1989; Schultz & Dickinson, 2000). A well-established measure of implicit learning is repetition priming, that is, the processing facilitation caused by the repeated encounter with a given stimulus. From the perspective of connectionism or predictive coding (Friston, 2005; McClelland, 1994; O'Reilly, Munakata, Frank, Hazy, & Contributors, 2012), such priming effects can be viewed as consequences of the continuous adaptation of the system aiming to reduce future prediction errors.

Thus, influences of the amount of associated semantic features on implicit word learning were investigated with a repetition priming design. Participants performed lexical decisions on 160 visual words differing in the amount of associated semantic features according to McRae et al. (2005), and 160 pseudowords. The complete stimulus set was presented twice; the lag between subsequent presentations of the same item varied randomly between 160 and 480 intermediate words. It is important to note that priming effects at such lags are assumed to be due to connection adaptations instead of residual activation (Becker, Moscovitch, Behrmann, & Joordens, 1997); semantic priming effects depending on residual activation typically disappear when prime and target are separated by several intervening items (e.g. Bentin & Feldman, 1990). Implicit learning was assessed by repetition priming effects on performance as well as N400 amplitudes which are typically reduced by repetition (Nagy & Rugg, 1989).

We found enhanced repetition priming for words with many as compared to few semantic features in both lexical decision accuracy and the N400 component. Thus, the richness of semantic representations advances repetition-induced changes in word processing considered to reflect implicit learning. These results seemed in line with feature-based connectionist network models that rely on a learning rule which is sensitive to semantic features, yielding substantial positive correlations between the number of features and the computed error driving connection adaptations (Cree, et al., 2006; Cree, et al., 1999; O'Connor, et al., 2009). Study 3 examined this correspondence more directly by simulating the obtained results with such a model, as further discussed in the general discussion section.

In sum, the results from study 2 suggest a novel and important impact of feature-based semantic richness on implicit learning and plasticity within the lexical conceptual system that should be taken into account when aiming to describe and understand reading development (Seidenberg & McClelland, 1989).

2.3 Linking N400 modulations to a model of semantic cognition (study 3)

An issue that was left unresolved in both studies 1 and 2 is the functional basis of the observed N400 modulations. Interestingly, an early comment by McClelland (1994; p. 61) suggested a relation between the N400 component and implicit prediction errors: “I do mean, though, that his or her cognitive system is in fact anticipating the future, and that a reaction can occur if these expectations are violated […]. They [such reactions] also, at least in language processing, generate large and robust evoked potentials, such as the N400 by Kutas & Hillyard (1980).” Indeed, it is interesting to note that N400 amplitudes seem to crucially depend on the fit between the information which is implicitly anticipated based on statistical regularities across levels of representations as represented in semantic memory (semantic context, relations between words, frequency of occurrence of single words…) and the actually observed information. As described in the introduction (section 1.1.1), implicit prediction error is implemented as network error (i.e. the difference between model-generated and correct output) in connectionist models, so that the proposal that N400 amplitudes reflect implicit prediction errors can be directly tested with a connectionist model of meaning. However, this early suggestion has not been further examined, and the recently published first implemented model of ERPs during reading did not consider different measures in the model, but instead directly started with the assumption that the N400 corresponds to the amount of semantic activation (Laszlo & Plaut, 2012).

Here, we tested the hypothesis that N400 amplitudes may reflect implicit prediction errors in the semantic system, represented by error values in a network model of meaning (McClelland, 1994). The model we used has successfully simulated a number of behavioral results in the semantic memory literature (Cree, et al., 2006; Mirman & Magnuson, 2008; O'Connor, et al., 2009), with 30 input units representing word form that map onto 2526 directly interconnected semantic feature units representing word meaning according to semantic feature production norms (McRae et al., 2005). To simulate the processing of word meaning, the corresponding word form was presented at the input layer and activation propagated to the semantic layer for 20 ticks (representing model time). The activation pattern produced at the semantic layer was interpreted as the activated word meaning. We were specifically interested in two measures: First, semantic network error, i.e. the difference between model-generated and correct activation across all semantic feature units, and second, the amount of activation across all semantic units which Laszlo & Plaut (2012) proposed to underlie N400 amplitudes. We simulated six N400 effects obtained in empirical research, including the results from studies 1 and 2 (Rabovsky, Sommer, & Abdel Rahman, 2012b;

Rabovsky, et al., 2012c), to examine the correspondence between N400 amplitude modulations and variations in both, semantic network error and semantic activation.

In line with our hypothesis, network error values were consistently in the same direction as N400 amplitudes. Like N400 amplitudes, error values were reduced for semantically related target words (simulation 1), while being enhanced for words with richer semantic representations (simulation 2; please also see study 1), and for low frequency words (simulation 3). Furthermore, error values decreased with repetition (simulation 4), and this repetition-induced decrease was stronger for words with richer semantic representations (simulation 5; please also see study 2), and for low frequency words (simulation 6). In contrast, contrary to the proposal by Laszlo & Plaut (2012), there was less correspondence between semantic activation and the N400. Like N400 amplitudes, activation was larger for words with richer semantic representations (simulation 2; see study 1). However, activation also increased with frequency, semantic priming, and repetition (simulations 1, 3, 4), and showed stronger repetition-induced *in*creases for words with richer semantic representations (simulation 5; see study 2) and low frequency words (simulation 6) which is all opposite to well-established N400 results (Kutas & Federmeier, 2011). Instead, the simulations seem better in line with the notion that semantic activation can improve lexical decision peformance, presumably by facilitating the crossing of decision thresholds when deciding that a stimulus is a word and not a pseudowords because words have meaning while pseudowords do not (Grondin, et al., 2009).

In sum, the results suggest a close relation between N400 amplitudes and semantic network error. Based on conceptualizing error values in connectionist models as implicit prediction error (McClelland, 1994; O'Reilly, et al., 2012; Rogers & McClelland, 2008), these results are taken to suggest that N400 amplitudes reflect implicit prediction error in the semantic system (McClelland, 1994).

3 General discussion and future directions

The present dissertation combined electrophysiology and connectionist modeling to explore influences of two measures of semantic richness in word reading and word learning, and to better understand the functional basis of the meaning-related N400 ERP component. Results showed no influences of the number of associates. In contrast, influences of feature-based semantic richness started at about 190 ms already, only about 20 to 30 ms after the activation of orthographic representations as indicated by N1 lexicality effects. Later on, words with many semantic features induced larger N400 amplitudes. Furthermore, the number of semantic features enhanced implicit word learning as assessed by repetition priming effects on lexical decision accuracy and N400 amplitudes. These results are in line with connectionist network models with feature-based representations of word meaning. Relating the observed results to such a model suggested the possibility that N400 amplitudes may reflect implicit prediction error in semantic memory while semantic activation may facilitate performance. I will first outline this proposal (section 3.1) to then examine how this perspective contributes to understanding semantic richness effects during reading (section 3.2) and influences of semantic richness on implicit word learning, while also suggesting a relation between the N400 component and implicit learning more generally (section 3.3). Open questions and suggestions for further research are discussed along the way.

3.1 The N400 as implicit prediction error in the semantic system

Simulations of the N400 effects obtained in study 1 and 2, as well as a number of well-established N400 effects (Kutas & Federmeier, 2011), consistently revealed a close correspondence between network error and N400 amplitudes while such a relation was not obtained for semantic activation values which seemed better in line with the notion that semantic activation can facilitate lexical decisions (Grondin, et al., 2009). These results do not fit with the recently proposed first model of ERPs during reading, which assumes a relation between the N400 and semantic activation (Laszlo & Plaut, 2012). Instead, the results are consistent with an early remark by McClelland (1994), who related N400 amplitudes to implicit prediction errors which are represented by error values in connectionist models (O'Reilly, et al., 2012; Rogers & McClelland, 2008).

In line with such an account in terms of implicit prediction error, N400 amplitudes seem to crucially depend on the similarity between actual observations and implicit anticipations based on represented occurrence probabilities as extracted from previously experienced regularities across levels of representation. From this perspective, N400 amplitudes are larger for semantic violations, low cloze probability sentence continuations or semantically unrelated targets, because words occur less frequently in the respective contexts and are therefore less expected, resulting in enhanced implicit prediction error. The represented occurrence probabilities for low frequency words are presumably generally rather low, resulting in enhanced implicit prediction error and thus enhanced N400 amplitudes for these words. Recent exposure to a word supposedly enhances its represented occurrence probability, giving rise to repetition-induced reductions of implicit prediction error and hence N400 amplitudes. In the following, I consider more specifically how this account explains and integrates the results obtained in the present ERP studies.

3.2 Semantic richness effects during reading from a network model perspective

Results from study 1 showed enhanced N400 amplitudes for words with many features. In line with this finding, simulation 2 in study 3 showed enhanced network error for these words. Conceptualizing network error as implicit prediction error (McClelland, 1994; O'Reilly, et al., 2012), this suggests higher implicit prediction error for words with many semantic features. While it seems pretty intuitive that implicit prediction error is higher for low frequency words, and for words that were not repeated or semantically primed, increased implicit prediction error for words with richer semantic representations may not seem very plausible at first sight. However, this relation might be explained by assuming that for every semantic feature, it is on average more probable that it is not involved in the currently relevant concept than that it is involved. This is definitely true for the features in the norms by McRae et al. (2005), and hence in the model, where on average each of the 2526 features occurs in only 2.87 out of the 541 concepts, so that the average probability for a feature to be involved in the representation of the current concept is below 1%. Thus, being involved is improbable, and even though the norms obviously do not cover the entire space of concepts and features, it seems reasonable to assume that this pattern generalizes beyond this reduced semantic space. Thus, every semantic feature may signal implicit prediction error when it is (unexpectedly) involved in the current concept, resulting in higher cumulative implicit prediction error and hence larger N400 amplitudes for words with more semantic features.

An important goal of study 1 was to pinpoint the moment when semantic richness effects first arise during reading. Interestingly, results showed that the influence of the number of features arose at 190 ms already, only about 20 to 30 ms after and temporally overlapping with influences of lexicality setting in during the left-lateralized N1 component which presumably reflects orthographic processing (McCandliss, et al., 2003). This fits well with fast access to semantics in reading (Dambacher, et al., 2009; Hauk, et al., 2006; Hoenig, et al., 2008; Kiefer, et al., 2008; Penolazzi, et al., 2007; Rabovsky, et al., 2012a; Skrandies, 1998) as well as partial information transmission and temporal overlap between sub-processes in reading, as implemented in connectionist models (Harm & Seidenberg, 2004). The mechanisms underlying this early feature effect are presently unclear – it might reflect an early start of the N400 effect, a possibility suggested by the similarity of the scalp distributions (please see Rabovsky, et al., 2012c), or alternatively a different preceding process, possibly at an interactive interface between visual word form processing and semantics. It would be intriguing to see whether this issue could be clarified by means of simulations. This seems difficult to accomplish with the model used herein, because the focus of the present simulations was on the N400 component which is related to semantic processes independent of perceptual modality and domain (Kutas & Federmeier, 2011). Thus, while quite an effort was made to approximate human semantic processing by implementing semantic representations based on empirical semantic feature production norms (McRae, et al., 2005), the input to semantics was highly oversimplified. However, a model with more elaborate visual input representations (features, letters, word forms…) to semantics that attempts to more fully capture the processes and interactions involved in visual word recognition might contribute to a clarification. Furthermore, in addition to the early and later feature effect, a more complete model may also allow for simulating the N1 lexicality effect, possibly as activation (or network error) in a visual word form layer. Establishing relations between ERP effects and such simulations would presumably also be informative for models of reading more generally (Harm & Seidenberg, 2004; Seidenberg & McClelland, 1989) which are often not specified in terms of timing.

The finding of ERP effects of the number of features but not the number of associates fits in with a recent behavioral study by Yap et al. (2011) reporting independent influences of the number of features but not associates when a number of lexical and semantic variables were controlled for. These results are in line with the present model as well as other models assuming feature-based representations of word meaning. In this context, it may be important to clarify that most feature-based models do not actually assume semantic representations to consist in lists of verbalizable features (Harm & Seidenberg, 2004; McRae, et al., 1997; Plaut & Shallice, 1993; Rogers & McClelland, 2008). Rather, as pointed out by McRae et al. (2005), semantic feature listings, as collected in feature production norms, are assumed to represent a temporary online abstraction of semantic representations, while the representations per se are assumed to be grounded in perceptual and motor systems according to repeated sensory experiences and interactions with the respective objects during concept acquisition, in line with proposals suggesting embodied semantic representations (Barsalou, 1999; Kiefer & Spitzer, 2001; Pulvermüller, 2005). Thus, verbalizable semantic features provide a helpful approximation to semantic representations, and the present data are in line with the notion that this approximation captures relevant aspects of meaning representation. However, before drawing general conclusions, one should consider that the present studies as well as the study by Yap et al. (2011) only used concrete words. It has been proposed that concrete and abstract words differ in their semantic representations, with concrete words relying more on semantic features and abstract words relying more on associations (Crutch & Warrington, 2005; Dunabeitia, Aviles, Afonso, Scheepers, & Carreiras, 2009). Accordingly, relevant semantic richness dimensions might not generalize across types of words. Indeed, a very recent study by Recchia and Jones (2012) found the number of semantic features to predict performance for concrete but not abstract words, while the number of semantic neighbors (words that occur in similar lexical contexts) facilitated performance for abstract words only. It would be interesting to test whether the pattern we found for the number of features and associates in concrete word processing might reverse similarly when using abstract words.

According to the proposal outlined above (section 3.1), enhanced activation for words with many features in the simulation of semantic richness effects during reading (study 3, simulation 2) predicts facilitated performance for these words, in line with previous evidence (Pexman, et al., 2008; Pexman, et al., 2002; Yap, et al., 2011). It seems interesting to note that this often observed semantic richness benefit in behavioral performance at some point seemed to be naturally explained by the enhancement of repetition effects for words with many semantic features obtained in study 2: “If the amount of semantic features associated with a given word enhances learning during every single encounter, repeated presentations should naturally entail the observed benefit" (Rabovsky, et al., 2012b, p. 1081). However, the simulations shed new light on this issue, suggesting that while the semantic richness benefit is indeed strengthened through enhanced repetition priming for words with many features (simulation 5; please also see the next section), it primarily relies on higher semantic activation for words with richer representations (please see simulation 2) which can facilitate crossing decision thresholds when discriminating between words and pseudowords (Grondin, et al., 2009).

In study 1, however, this often observed benefit was not obtained. The lack of the behavioral effect was most probably due to the orthographically rather untypical pseudowords, which had significantly lower bigram and trigram frequencies as compared to the words, so that semantic processes presumably contributed little to lexical decision performance, because decisions could instead be based on lower orthographic levels of representation. This explanation receives support from the fact that regression analyses of response times to the same stimuli as retrieved from the English Lexicon Project (ELP; Balota et al., 2007) showed the expected facilitative influence of the number of features but not associates (please see Rabovsky, et al., 2012c, p. 7). This crucial effect of decision threshold and how it can be influenced by all sorts of context variables independent of the manipulation of interest highlights the usefulness of ERPs in investigating internal processes beyond the presently task-relevant aspects, and serving to inform models of semantic cognition.

3.3 Semantic richness and implicit learning from a network model perspective

In study 2, feature-based semantic richness was found to enhance implicit word learning as reflected in repetition-induced N400 amplitude reductions and increases in lexical decision accuracy. In line with these results, repetition resulted in enhanced reductions of network error and enhanced increases of semantic activation for words with many as compared to few semantic features in study 3 (simulation 5). When taken together with the finding of larger N400 amplitudes (and enhanced network error) for words with many features during single word presentation (study 1; study 3, simulation 2), and the model-based interpretation of the N400 as implicit semantic prediction error (study 3), this enhanced repetition effect (presumably reflecting enhanced connection adaptation) can be naturally explained by the widely shared assumption that learning is based on prediction error (den Ouden, et al., 2009; Friston, 2005, 2009; McClelland, 1994; McLaren, 1989; Rogers & McClelland, 2008; Schultz & Dickinson, 2000; Tobler, et al., 2006). A similar pattern as apparent in the joint consideration of studies 1 and 2, namely that the condition with enhanced N400 amplitudes during single word presentation shows enhanced repetition effects in performance and ERPs, can also be found for low frequency words (Rugg, 1990; please see simulations 3 and 6). It seems interesting to note that this relation between N400 amplitudes during initial presentation and repetition effects does not depend on performance during single word processing: While performance is generally better for words with richer semantic representations, the opposite is true for low frequency words.

More generally, if the N400 reflects implicit prediction error, and implicit prediction error drives connection adaptation, as often assumed (e.g. McClelland, 1994; McLaren, 1989; Rogers & McClelland, 2008; Schultz & Dickinson, 2000; Tobler, et al., 2006), then enhanced N400 amplitudes should entail enhanced connection adaptation, i.e. enhanced implicit memory formation, as for example reflected in stronger repetition effects. There are indeed some pieces of evidence supporting this suggestion, most notably a study by Schott, Richardson-Klavehn, Heinze, & Düzel (2002) reporting larger N400 amplitudes during a learning phase to predict implicit memory (as assessed by repetition priming effects on stem completion in the absence of explicit memory) during test. However, further research on the relation between N400 amplitudes and implicit memory formation seems highly desirable. In this context, it seems important to note that the intrinsic theoretical relation between prediction error and connection adaptation makes it presently difficult to unequivocally decide whether N400 amplitudes reflect implicit prediction error or whether they might rather reflect the connection adaptations driven by implicit prediction error. Further research is needed to clearly disentangle these possibilities, possibly by analyzing N400 data in the frame of the neurobiologically detailed predictive coding model suggested by Friston and colleagues (Friston, 2005, 2009; Garrido, Kilner, Stephan, & Friston, 2009).

Finally, the reported driving influence of semantic richness on language learning seems to suggest some intriguing implications to be addressed, for instance, concerning reading difficulties and second language acquisition. Furthermore, it remains to be explored in how far the reported driving influence of semantic richness on implicit visual word learning generalizes to other perceptual modalities (e.g. visual vs. auditory) and domains (e.g. language vs. object recognition). The present model-based account would suggest a generalization, because the driving mechanism is functionally localized in the semantic layer, independent of the somewhat arbitrary input. However, further research is required to examine this suggestion.

3.4 Conclusions

In sum, the present work combined ERPs and a connectionist network model to explore independent influences of the number of semantic features and associates during word reading and word learning, and to better understand the computational mechanisms underlying the N400 ERP component. The results support feature-based network models of semantic cognition, with fast initial access to semantic features during reading, arising within the first 200 ms and resuming during the N400 segment, as well as a driving influence of feature-based semantic richness on implicit word learning. Simulations using a network model with empirically derived semantic features suggest that semantic activation can facilitate lexical decisions, while network error closely corresponds to N400 amplitudes. Based on conceptualizing network error as implicit prediction error (McClelland, 1994; O'Reilly, et al., 2012), these results are taken to suggest that N400 amplitudes may reflect implicit prediction error in semantic memory

References

Abdel Rahman, R., & Sommer, W. (2008). Seeing what we know and understand: How knowledge shapes perception. *Psychonomic Bulletin & Review, 15*(6), 1055-1063. doi: 10.3758/Pbr.15.6.1055

Adelman, J. S., Brown, G. D., & Quesada, J. F. (2006). Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychological Science, 17*(9), 814-823. doi: 10.1111/j.1467-9280.2006.01787.x

Amsel, B. D. (2011). Tracking real-time neural activation of conceptual knowledge using single-trial event-related potentials. *Neuropsychologia, 49*(5), 970-983. doi: 10.1016/j.neuropsychologia.2011.01.003

Andrews, M., Vigliocco, G., & Vinson, D. (2009). Integrating experiential and distributional data to learn semantic representations. *Psychological Review, 116*(3), 463-498. doi: 10.1037/A0016261

Baggio, G., & Hagoort, P. (2011). The balance between memory and unification in semantics: A dynamic account of the N400. *Language and Cognitive Processes, 26*(9), 1338-1367. doi: 10.1080/01690965.2010.542671

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39*(3), 445-459.

Barrett, S. E., & Rugg, M. D. (1989). Event-related potentials and the semantic matching of faces. *Neuropsychologia, 27*(7), 913-922.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(4), 577-609; discussion 610-560.

Becker, S., Moscovitch, M., Behrmann, M., & Joordens, S. (1997). Long-term semantic priming: A computational account and empirical evidence. *Journal of Experimental Psychology: Learning Memory and Cognition, 23*(5), 1059-1082.

Bentin, S., & Feldman, L. B. (1990). The contribution of morphological and semantic relatedness to repetition priming at short and long lags: Evidence from Hebrew. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 42*(4), 693-711.

Bentin, S., Mccarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology, 60*(4), 343-355.

Buchanan, L., Westbury, C., & Burgess, C. (2001). Characterizing semantic space: Neighborhood effects in word recognition. *Psychonomic Bulletin & Review, 8*(3), 531-544.

Collins, A. M., & Loftus, E. F. (1975). Spreading activation theory of semantic processing. *Psychological Review, 82*(6), 407-428.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8*(2), 240-247.

Cree, G. S., McNorgan, C., & McRae, K. (2006). Distinctive features hold a privileged status in the computation of word meaning: Implications for theories of semantic memory. *Journal of Experimental Psychology: Learning Memory and Cognition, 32*(4), 643-658. doi: 10.1037/0278-7393.32.4.643

Cree, G. S., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science, 23*(3), 371-414.

Crutch, S. J., & Warrington, E. K. (2005). Abstract and concrete concepts have structurally different representational frameworks. *Brain, 128*, 615-627. doi: 10.1093/brain/awh349

Dambacher, M., Rolfs, M., Gollner, K., Kliegl, R., & Jacobs, A. M. (2009). Event-related potentials reveal rapid verification of predicted visual input. *Plos One, 4*(3). doi: 10.1371/journal.pone.0005047

Debruille, J. B. (2007). The N400 potential could index a semantic inhibition. *Brain Research Reviews, 56*(2), 472-477. doi: 10.1016/j.brainresrev.2007.10.001

den Ouden, H. E. M., Daunizeau, J., Roiser, J., Friston, K. J., & Stephan, K. E. (2010). Striatal prediction error modulates cortical coupling. *Journal of Neuroscience, 30*(9), 3210-3219. doi: 10.1523/Jneurosci.4458-09.2010

den Ouden, H. E. M., Friston, K. J., Daw, N. D., McIntosh, A. R., & Stephan, K. E. (2009). A dual role for prediction error in associative learning. *Cerebral Cortex, 19*(5), 1175-1185. doi: 10.1093/cercor/bhn161

Dunabeitia, J. A., Aviles, A., Afonso, O., Scheepers, C., & Carreiras, M. (2009). Qualitative differences in the representation of abstract versus concrete words: Evidence from the visual-world paradigm. *Cognition, 110*(2), 284-292. doi: 10.1016/j.cognition.2008.11.012

Dunabeitia, J. A., Aviles, A., & Carreiras, M. (2008). NoA's ark: Influence of the number of associates in visual word recognition. *Psychonomic Bulletin & Review, 15*(6), 1072-1077. doi: 10.3758/Pbr.15.6.1072

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language, 41*(4), 469-495.

Federmeier, K. D., & Laszlo, S. (2009). Time for meaning: Electrophysiology provides insights into the dynamics of representation and processing in semantic memory. *Psychology of Learning and Motivation: Advances in Research and Theory, Vol 51, 51*, 1-44. doi: 10.1016/S0079-7421(09)51001-8

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B-Biological Sciences, 360*(1456), 815-836. doi: 10.1098/rstb.2005.1622

Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences, 13*(7), 293-301. doi: 10.1016/j.tics.2009.04.005

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology, 120*(3), 453-463. doi: 10.1016/j.clinph.2008.11.029

Gopnik, A., & Wellmann, H. M. (1994). The theory theory. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture*: Cambridge University Press.

Grondin, R., Lupker, S. J., & McRae, K. (2009). Shared features dominate semantic richness effects for concrete concepts. *Journal of Memory and Language, 60*(1), 1-19. doi: 10.1016/j.jml.2008.09.001

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304*(5669), 438-441. doi: 10.1126/science.1095455

Hargreaves, I. S., Pexman, P. M., Johnson, J. C., & Zdrazilova, L. (2012). Richer concepts are better remembered: number of features effects in free recall. *Frontiers in Human Neuroscience, 6*. doi: 10.3389/fnhum.2012.00073

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychological Review, 111*(3), 662-720. doi: 10.1037/0033-295X.111.3.662

Hauk, O., Davis, M. H., Ford, M., Pulvermuller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage, 30*(4), 1383-1400. doi: 10.1016/j.neuroimage.2005.11.048

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. *Journal of Cognitive Neuroscience, 20*(10), 1799-1814. doi: 10.1162/jocn.2008.20123

Holcomb, P. J., Kounios, J., Anderson, J. E., & West, W. C. (1999). Dual-coding, context-availability, and concreteness effects in sentence comprehension: An electrophysiological investigation. *Journal of Experimental Psychology: Learning Memory and Cognition, 25*(3), 721-742.

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review, 10*(4), 785-813.

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. *Journal of Neuroscience, 28*(47), 12224-12230. doi: 10.1523/Jneurosci.3579-08.2008

Kiefer, M., & Spitzer, M. (2001). The limits of a distributed account of conceptual knowledge. *Trends in Cognitive Sciences, 5*(11), 469-471. doi: 10.1016/S1364-6613(00)01798-8

Kounios, J., Green, D. L., Payne, L., Fleck, J. I., Grondin, R., & Mcrae, K. (2009). Semantic richness and the activation of concepts in semantic memory: Evidence from event-related potentials. *Brain Research, 1282*, 95-102. doi: 10.1016/j.brainres.2009.05.092

Kounios, J., & Holcomb, P. J. (1994). Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20*(4), 804-823.

Kutas, M. (1993). In the company of other words: Electrophysiological evidence for single-word and sentence context effects. *Language and Cognitive Processes, 8*(4), 533-572.

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

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621-647. doi: 10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203-205.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*(5947), 161-163.

Laszlo, S., & Plaut, D. C. (2012). A neurally plausible parallel distributed processing model of event-related potential word reading data. *Brain and Language, 120*(3), 271-281. doi: 10.1016/j.bandl.2011.09.001

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience, 9*(12), 920-933. doi: 10.1038/nrn2532

Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review, 7*(4), 618-630.

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences, 7*(7), 293-299. doi: 10.1016/S1364-6613(03)00134-7

McClelland, J. L. (1994). The interaction of nature and nurture in development: A parallel distributed processing perspective. In P. E. P. Bertelson, G. d'Ydewalle (Ed.), *International Perspectives on Psychological Science* (Vol. 1). United Kingdom: Erlbaum.

McLaren, I. (1989). The computational unit as an assembly of neurons: An implementation of an error correcting learning algorithm. In R. Durbin, Miall, C., Mitchison, G. (Ed.), *The computing neuron* (pp. 160-178). Amsterdam, the Netherlands: Addison-Wesley.

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods, 37*(4), 547-559.

McRae, K., deSa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General, 126*(2), 99-130.

Mirman, D., & Magnuson, J. S. (2008). Attractor dynamics and semantic neighborhood density: Processing is slowed by near neighbors and speeded by distant neighbors. *Journal of Experimental Psycholy: Learning, Memory, and Cognition, 34*(1), 65-79. doi: 10.1037/0278-7393.34.1.65

Müller, O., Dunabeitia, J. A., & Carreiras, M. (2010). Orthographic and associative neighborhood density effects: What is shared, what is different? *Psychophysiology, 47*(3), 455-466. doi: 10.1111/j.1469-8986.2009.00960.x

Nagy, M. E., & Rugg, M. D. (1989). Modulation of event-related potentials by word repetition: The effects of inter-item lag. *Psychophysiology, 26*(4), 431-436.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods Instruments & Computers, 36*(3), 402-407.

O'Connor, C. M., Cree, G. S., & McRae, K. (2009). Conceptual hierarchies in a flat attractor network: Dynamics of learning and computations. *Cognitive Science, 33*(4), 665-708. doi: 10.1111/j.1551-6709.2009.01024.x

O'Reilly, R. C., Munakata, Y., Frank, M. J., Hazy, T. E., & Contributors. (2012). *Computational Cognitive Neuroscience. Wiki Book* (1 ed.). URL: http://ccnbook.colorado.edu.

Penolazzi, B., Hauk, O., & Pulvermüller, F. (2007). Early semantic context integration and lexical access as revealed by event-related brain potentials. *Biological Psychology, 74*(3), 374-388. doi: 10.1016/j.biopsycho.2006.09.008

Pexman, P. M., Hargreaves, I. S., Edwards, J. D., Henry, L. C., & Goodyear, B. G. (2007). The neural consequences of semantic richness: When more comes to mind, less activation is observed. *Psychological Science, 18*(5), 401-406.

Pexman, P. M., Hargreaves, I. S., Siakaluk, P. D., Bodner, G. E., & Pope, J. (2008). There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition. *Psychonomic Bulletin & Review, 15*(1), 161-167. doi: 10.3758/Pbr.15.1.161

Pexman, P. M., Holyk, G. G., & Monfils, M. H. (2003). Number-of-features effects and semantic processing. *Memory & Cognition, 31*(6), 842-855.

Pexman, P. M., Lupker, S. J., & Hino, Y. (2002). The impact of feedback semantics in visual word recognition: Number-of-features effects in lexical decision and naming tasks. *Psychonomic Bulletin & Review, 9*(3), 542-549.

Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology, 10*(5), 377-500.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience, 6*(7), 576-582. doi: 10.1038/nrn1706

Rabovsky, M., Sommer, W., & Abdel Rahman, R. (2012a). Depth of conceptual knowledge modulates visual processes during word reading. *Journal of Cognitive Neuroscience, 24*(4), 990-1005.

Rabovsky, M., Sommer, W., & Abdel Rahman, R. (2012b). Implicit word learning benefits from semantic richness: Electrophysiological and behavioral evidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(4), 1076-1083. doi: 10.1037/a0025646

Rabovsky, M., Sommer, W., & Abdel Rahman, R. (2012c). The time course of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience, 6*. doi: 10.3389/fnhum.2012.00011

Recchia, G., & Jones, M. N. (2012). The semantic richness of abstract concepts. *Frontiers in Human Neuroscience, 6*. doi: 10.3389/fnhum.2012.00315

Rogers, T. T., & McClelland, J. L. (2008). Precis of semantic cognition: A parallel distributed processing approach. *Behavioral and Brain Sciences, 31*(6), 689-749. doi: 10.1017/S0140525x0800589x

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*(4), 573-605.

Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high-frequency and low-frequency words. *Memory & Cognition, 18*(4), 367-379.

Schott, B., Richardson-Klavehn, A., Heinze, H. J., & Düzel, E. (2002). Perceptual priming versus explicit memory: Dissociable neural correlates at encoding. *Journal of Cognitive Neuroscience, 14*(4), 578-592.

Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience, 23*, 473-500.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*(4), 523-568.

Skrandies, W. (1998). Evoked potential correlates of semantic meaning - A brain mapping study. *Cognitive Brain Research, 6*(3), 173-183.

Tobler, P. N., O'Doherty, J. P., Dolan, R. J., & Schultz, W. (2006). Human neural learning depends on reward prediction errors in the blocking paradigm. *Journal of Neurophysiology, 95*(1), 301-310.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of Memory* (pp. 381-403). New York: Academic Press.

Van Petten, C., & Rheinfelder, H. (1995). Conceptual relationships between spoken words and environmental sounds: Event-related brain potential measures. *Neuropsychologia, 33*(4), 485-508.

West, W. C., & Holcomb, P. J. (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: An electrophysiological investigation. *Journal of Cognitive Neuroscience, 12*(6), 1024-1037.

Yap, M. J., Pexman, P. M., Wellsby, M., Hargreaves, I. S., & Huff, M. J. (2012). An abundance of riches: cross-task comparisons of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience, 6*. doi: 10.3389/fnhum.2012.00072

Yap, M. J., Tan, S. E., Pexman, P. M., & Hargreaves, I. S. (2011). Is more always better? Effects of semantic richness on lexical decision, speeded pronunciation, and semantic classification. *Psychonomic Bulletin & Review, 18*(4), 742-750. doi: 10.3758/s13423-011-0092-y

Yee, E., Overton, E., & Thompson-Schill, S. L. (2009). Looking for meaning: Eye movements are sensitive to overlapping semantic features, not association. *Psychonomic Bulletin & Review, 16*(5), 869-874. doi: 10.3758/Pbr.16.5.869

Original articles

I Rabovsky, M., Sommer, W., & Abdel Rahman, R. (2012). The time course of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience, 6:11*. doi: 10.3389/fnhum.2012.00011

II Rabovsky, M., Sommer, W., & Abdel Rahman, R. (2012). Implicit word learning benefits from semantic richness: Electrophysiological and behavioral evidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(4), 1076-1083. doi: 10.1037/a0025646

III Rabovsky, M. & McRae, K. (under review). Simulating the N400 ERP component as semantic network error: Insights from a feature-based connectionist attractor model of word meaning. *Manuscript under review for Cognition.*

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt,

1. dass ich die vorliegende Arbeit selbständig und ohne unerlaubte Hilfe verfasst habe,
2. dass ich mich nicht anderwärts um einen Doktorgrad beworben habe und noch keinen Doktorgrad der Psychologie besitze,
3. dass mir die zugrunde liegende Promotionsordnung vom 3. August 2006 bekannt ist.

Berlin, den

Milena Rabovsky