ABSTRACT

Title of Document:	THE PREDICTIVE NATURE OF LANGUAGE COMPREHENSION
	Ellen Frances Lau
	Doctor of Philosophy 2009
Directed By:	Professor Colin Phillips Department of Linguistics

This dissertation explores the hypothesis that predictive processing—the access and construction of internal representations in advance of the external input that supports them—plays a central role in language comprehension. Linguistic input is frequently noisy, variable, and rapid, but it is also subject to numerous constraints. Predictive processing could be a particularly useful approach in language comprehension, as predictions based on the constraints imposed by the prior context could allow computation to be speeded and noisy input to be disambiguated.

Decades of previous research have demonstrated that the broader sentence context has an effect on how new input is processed, but less progress has been made in determining the mechanisms underlying such contextual effects. This dissertation is aimed at advancing this second goal, by using both behavioral and neurophysiological methods to motivate predictive or top-down interpretations of contextual effects and to test particular hypotheses about the nature of the predictive mechanisms in question.

The first part of the dissertation focuses on the lexical-semantic predictions made possible by word and sentence contexts. MEG and fMRI experiments, in conjunction with a meta-analysis of the previous neuroimaging literature, support the claim that an ERP effect classically observed in response to contextual manipulations—the N400 effect—reflects facilitation in processing due to lexical-semantic predictions, and that these predictions are realized at least in part through top-down changes in activity in left posterior middle temporal cortex, the cortical region thought to represent lexical-semantic information in long-term memory,. The second part of the dissertation focuses on syntactic predictions. ERP and reaction time data suggest that the syntactic requirements of the prior context impacts processing of the current input very early, and that predicting the syntactic position in which the requirements can be fulfilled may allow the processor to avoid a retrieval mechanism that is prone to similarity-based interference errors.

In sum, the results described here are consistent with the hypothesis that a significant amount of language comprehension takes place in advance of the external input, and suggest future avenues of investigation towards understanding the mechanisms that make this possible.

THE PREDICTIVE NATURE OF LANGUAGE COMPREHENSION

By

Ellen Frances Lau

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2009

Advisory Committee:
Professor Colin Phillips, Chair
Professor David Poeppel
Professor Norbert Hornstein
Associate Professor William Idsardi
Associate Professor Michael Dougherty

© Copyright by
Ellen Frances Lau
2009

Dedication

for John, Sarah, Rachel, and Thomaz

Acknowledgements

I'm not going to be especially short here. I'm not really thanking everyone for making this particular dissertation possible, because I think that would be kind of silly. I am thanking you for everything you gave me across these years of grad school that made my life richer.

It is impossible to sum up or quantify what I've gained from working with Colin over the years. I don't think anybody else would have pushed me so hard or gotten so much out of me, and I can’t imagine a more dedicated advisor—Colin always read everything, no matter how last minute I sent it, met whenever I wanted, gave me space when I needed it, and was unfailingly supportive when I hit tough spots. Even when I was doing something that he didn't think was worthwhile, if my heart was really set on it he'd usually let me go ahead and do it anyway. Sometimes I'd be right, more often I'd be wrong, but altogether I learned a lot, and I think there just aren't that many advisors that let you go on and make your own mistakes that way. So all in all, I feel really lucky to have been his student.

Before I knew David I wouldn't have thought that you could play the game to the max and still be a truly good person. Even though I don't think I'll ever be such a good player, that always puts a smile on my face. I learned a ton about everything from David and since he’s always trying to do a million different things at once, I'm really grateful for the time and resources he put into helping me. And both him and Norbert saved me a bunch of times by reminding me that when things get too complicated, at some point you've just got to pretend you're a rock and go for the dumb story. That's been a really good lesson.

I could say a million things about Norbert, but the bottom line is that he gave me, as he gives everyone, not only cookies, but boundless support that gave me the energy to keep going and to be a better colleague to everyone around me. Even though we both agree now that you can't really teach anybody the wisdom you have accumulated, I can say that it seems like the longer I know him the fewer dumb things I end up doing.

I don't know why these three guys still kept believing in me at some points, but I will never forget it and I am very, very grateful to them for it.

Bill has been a wonderful source of insight and support throughout graduate school and I will really miss being able to pop into his office when I need a fresh perspective, on any topic.

Fernanda Ferreira gave me the best start in research that you possibly could. She trained me, trusted me, supported me, and gave me the space to explore what I was interested in, and she did so much to give me a good start in graduate school. She was also a good friend and source of support during a time when I was pretty unsure of myself. I am eternally grateful to her.

Alan Munn and Cristina Schmitt introduced me to generative grammar just when I needed it, and just gave an incredible amount of support to me as an undergraduate. Although I don't think I'll ever be as head-over-heels for formal syntax as I was when I was 20, they changed the direction of my research path unalterably.

Thank you to Jim Zacks, who gave me good practical advice when I was deciding what to do with my life, and Rick DeShon, who taught me stats.

Thank you to Al Braun, who gave me amazing opportunities that I wouldn't have had otherwise.

Thank you to Nina Kazanina and Masaya Yoshida, who took me under their wing when I was first at Maryland, gave me great examples of how to have fun and do good work in graduate school, and kept an eye on me even after they left.

Thank you to Ming Xiang, who's always one step ahead of me, showing me the way, and who always warns me by tipping her head to the side before she asks me something that will force me to re-examine all my assumptions.

Thank you to all my classmates, who made our five years together so fun and stimulating: Yuval Marton, Rebecca McKeown, Chizuru Nakao, Eri Takahashi, Stacey Trock, and especially Phil Monahan, who I puzzled through analysis-by-synthesis and my first real teaching experience with.

Thank you to my 'bigger brother and sister' classmates: Clare Stroud, Matt Wagers, Jon Sprouse, and Diogo Almeida (who actually is my big brother now!), who arrived at Maryland when I did and who I've been through so much with. I never would have made it without you guys.

Thank you to Henny Yeung, who always pushes me to reassess everything and who's been a great friend through these years despite the distance.

Thank you to Tim Hunter, who really saved me in the later stages of this dissertation by always being willing to sit down and help me make sure I wasn't saying anything too stupid about the formal stuff and by reminding me that there's no magic in parsing.

Thank you to Suzanne Dikker and Adrian Staub, for several enlightening discussions about prediction in language over the years.

Thank you to Akira Omaki, who gave me lots of syntax and parsing advice and told me the secret to getting past writer's block.

Thank you to Jeff Walker, who brightened everybody's life who worked with MEG, the best lab manager I've ever seen.

Thank you to Alexander Williams for helping me hide my unregistered car as I finished writing this dissertation and for reminding me why semantics is interesting.

Thanks so much to Kathi Faulkingham who always stopped what she was doing to help me and who found my keys for me numerous times over the years.

Thanks to Kim Kwok, for lots of help with bureaucracy and for rescuing me from a wardrobe malfunction on one of the critical last days of dissertation-writing.

Thanks to Annie Gagliardi and Dave Kush for taking me up to the last part of Marie Mount Hall that I had not explored over my six years here.

Thanks to So One Hwang and Ariane Rhone, for the lasagna and for putting up with me whenever I came into distract them.

Thanks to Jaiva Larsen, Maria Sanchez, Bernie Wandel, and the All Souls Church Unitarian for keeping my spirits up as I was getting older and some parts of life were getting harder.

Thanks to Sampson, who gave me a lot of unconditional love at the hardest parts, the beginning and the end.

Thanks to the lady at the College Park UPS for helping me and Thomaz maintain our long-distance relationship for a year and a half by giving me lots of patient advice on the cheapest way to send baked goods to Brazil.

For support, help, and friendship along the way, thanks to Andrea Zukowski, Brian Dillon, Peggy Antonisse, Tonia Bleam, Jeff Lidz, Rob Magee, Alda Ottley, Nuria Abdulsabur, Mike Shvartsman, Elika Bergelson, Chris Dyer, Johannes Jurka, Meredith Reitz, Ryu Hashimoto, Amy Weinberg, Jonathan Simon, Matthias Scharinger, Veronica Figueroa, Julian Jenkins, Dave Huber, Moti Lieberman, Ching-fen Hsu, Walter Carr, Bruce Swett, Kiel Christianson, Minna Lehtonen, Pedro Alcocer, Shannon Hoerner, Xing Tian, Alex Drummond, Rob Fiorentino, Hajime Ono, Cilene Rodrigues, Andrew Nevins, Ilhan Cagri, Scott Fults, Katya Rozanova.

I feel very, very lucky to be married to someone who supports my career so much. Even though it makes our life difficult in a lot of ways, Thomaz has always encouraged me to keep going, even when *I* have doubted what I was doing. And through his positive example, he just made it possible for me to keep doing this kind of work long-term. When I started doing this, I was succeeding all right on paper, but in a way that was really emotionally unsustainable. I don't know if I would have made it through the PhD if Thomaz hadn't taught me how to keep doing what I love in a better way.

Finally, I'm just so lucky to have the family I have--both the old one and the new one--that there's nothing more I can say about it.

Table of Contents

List of Tables

List of Figures

1 Introduction

'the brain does not depend on continuous input from the external world to generate perceptions, but only to modulate them contextually…'
-Llinas, 2001

1.1 Overview

The goal of this dissertation is to both provide stronger evidence for top-down and predictive mechanisms in language comprehension and to outline a framework for studying these mechanisms more systematically. Key aims of psycholinguistics are to determine the sequence of mental operations underlying (1) access of abstract lexical representations from the input and (2) composition of these individual representations into higher-level structure. The hypothesis I explore in this dissertation is that in both cases the sequence of operations often begins before all of the external input supporting these representations is presented, based on knowledge of dependencies—both discrete and probabilistic—that hold between the prior context and the upcoming input.

Many cognitive science researchers now assume the existence of 'top-down' mechanisms—mechanisms through which representations of the broader context influence how the current input is represented—because of the many studies that have demonstrated contextual effects on the processing of words or phonemes during comprehension (Ladefoged & Broadbent, 1957; Warren, 1970; Fischler & Bloom, 1979; Ganong, 1980; Stanovich & West, 1981; Elman & McClelland, 1988; Duffy, Henderson, & Morris, 1989). However, it has proven surprisingly difficult to find unambiguous evidence for top-down effects in language processing, and it has proved equally difficult to advance beyond the claim that top-down mechanisms exist to more specific hypotheses about how these mechanisms are implemented. In the work described here I consider the factors that have led to these difficulties and attempt to overcome them in a series of studies examining top-down predictive mechanisms in lexical access and syntactic structure-building with reaction time measures, ERP, MEG, and fMRI.

While much of the research and theory on top-down and predictive effects in recent years has focused on vision (e.g. Ullman, 1984; Knill & Richards, 1996; Allbright & Stoner, 2002; Rao, Olshausen, & Lewicki, 2002; Lee & Mumford, 2003; Murray, Schrater & Kersten, 2004; Hawkins & Blakeslee, 2004; Yuille & Kersten, 2006), prediction is particularly likely to be both useful and easy to study in the domain of language, because the prior input imposes multiple constraints and likelihoods on the upcoming input, many of which have already been well-characterized by linguistic theory. Furthermore, prediction is a potential solution to some of the most difficult problems faced in language comprehension: linguistic input, particularly in the auditory modality, is frequently obscured by noise and is highly variable across speakers and contexts, and the rapid rate of speech and normal reading (200-300 ms/word) makes only a small window of time available for the multiple stages of processing required for each word (i.e., preprocessing by visual or auditory systems, access of the stored lexico-semantic representation, and integration into the syntactic, semantic, and discourse representations currently under construction). Top-down mechanisms may solve these problems by speeding computation on incoming words through pre-activation and pre-processing, facilitating disambiguation of the upcoming signal, and limiting the need for 'backward-oriented' memory retrieval processes that are prone to errors of interference.

This dissertation joins other recent work (Federmeier & Kutas, 1999; Wicha, Morena, & Kutas, 2004; Van Wassenhove, Grant, & Poeppel, 2005; Delong, Urbach, & Kutas, 2005; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; Staub & Clifton, 2006; Dikker, Rabagliati, & Pylkkänen, 2009) in aiming to provide stronger evidence for the existence of top-down mechanisms in language comprehension and to develop better models of how these mechanisms interact with bottom-up information processing.

1.2 Forms of prediction

It is a standard assumption that language processing, like most other cognitive tasks, involves manipulating two kinds of representations: 'stored' long-term memory representations such as lexical representations, and 'derived' representations that are built on the fly from individual stored representations according to stored rules of combination and only temporarily maintained in some kind of working memory, such as syntactic or thematic representations of sentences or mental models of relevant entities in a discourse. Predictive mechanisms may operate over either of these representational types, or both.

Perhaps the more intuitive form of prediction is activation of stored representations prior to the external stimulus. On the assumption that recognition of a meaningful stimulus involves some kind of change in activity in the neurons that represent this stimulus in the brain—a change that I will refer to as 'activation' for ease of presentation—a predictive mechanism could allow neurons to be activated prior to the external stimulus by a context that predicted that stimulus, rather than only by connections from lower-level perceptual areas based on the bottom-up input. How exactly this would lead to faster and more accurate recognition is highly dependent on the mechanisms that are assumed for bottom-up recognition, but one example is a threshold system in which less bottom-up processing is required to reach the required threshold of evidence for deciding on a particular candidate when activation prior to the stimulus has already provided some evidence. I will refer to this kind of prediction as pre-activation.

A second form of prediction is pre-construction or pre-updating of the derived structured representations. In many cases, recognition of a particular stimulus serves as the input to a process with a higher goal, such as mapping out a trajectory of movement towards a preferred object or understanding a propositional statement based on the individual words that compose it. If the location of the object or properties of the word can be predicted before the stimulus is presented, then work on these higher representations could also proceed prior to the stimulus—the trajectory of the hand could begin to be planned, an entire noun phrase could be constructed on the basis of a determiner, or the predicted semantic[1] interpretation could begin to be compared against world knowledge. Not only would such pre-construction make processing more efficient, but it could also avoid the memory costs that might be incurred by waiting until later to construct the higher representation. In Chapters 5 and 6, I will discuss cases in which syntactic dependencies might be predictively constructed.

One of the goals of this thesis is to show not only that predictive mechanisms play a role in language processing, but also to consider which of these forms of prediction the evidence points to in particular cases. In Chapters 2 and 3, I describe evidence that suggests that lexical processing in context involves predictive pre-activation of stored lexical representations. In Chapters 5 and 6, I provide evidence for syntactic prediction,

---

[1] In keeping with practice in much of the psycholinguistic literature, I will use the term ‘semantic’ very broadly to indicate any aspect of relating linguistic symbols to meaning.

and suggest that the effects observed may be due to pre-construction of syntactic structure ahead of bottom-up input.

1.3 Defining prediction

In this dissertation I will be assuming a critical distinction between the existence of a statistical dependency between the context and the upcoming input and the system's use of this dependency during processing. In other words, although a context may 'predict' a subsequent stimulus, in the sense that the context is very likely to be followed by a particular stimulus, that statistical generalization may or may not be used in processing. Even when we have reason to think that the system has knowledge of a dependency between context and the upcoming input—as in the case of grammatical requirements—we still need empirical evidence to determine whether this knowledge is being used predictively, that is, ahead of the input.

I also want to draw attention to a subtler distinction between the more general notion of 'top-down' processes in perception and the potentially narrower notion of 'prediction'. Top-down processing generally refers to cases in which activity at a 'higher' level of representation (more abstract and/or spanning more input) affects activity at a lower level of representation. For example, a top-down effect of sentence context on lexical access might be one in which activity at the level at which lexical representations are stored is dependent on activity at the level at which derived representations like the partial syntactic or thematic structure are maintained.

For me, predictive mechanisms are those that involve constructing representations on the basis of context and without the benefit of external information, and this may often be realized through top-down information flow, as in the example above. However, facilitation due to higher-level information is not limited to prediction, and not all top-down mechanisms assume that processing takes place prior to the external input. Top-down effects can also be realized non-predictively, through mechanisms that use the context as filters on representations activated by external input. For example, top-down effects can include cases where the context is used to help decide which of several candidate representations that are highly activated by the bottom-up input should be selected for further processing, if the context alters the activity at the level at which the candidate representations are stored. If we consider predictive effects to be restricted to situations in which the system takes some steps towards constructing a representation that will only be confirmed by bottom-up input in the future, such cases would not qualify.

This distinction is important to keep in mind because several theories of top-down processing imply that context only begins to have an effect on perception after a rough first-pass 'gist' representation of the bottom-up input makes it up to the top of the processing hierarchy (e.g., Hochstein & Ahissar, 2002; Bar, 2007). One theory that is particularly explicit about this is the version of the cohort model of word recognition presented in Marslen-Wilson (1987), which assigns a major role to top-down effects, but importantly, argues against actual prediction:

> A lexical unit is assumed to become active when the sensory input matches the acoustic-phonetic pattern specified for that unit. The model prohibits top-down activation of these units in normal word-recognition, so that *only* the sensory input can activate a unit. There is no contextually driven pre-selection of candidates, so that words cannot become active as potential percepts without some bottom-up (sensory) input to the structures representing these words. … Once the word-initial cohort has been accessed, and the model has entered into the selection phase, then top-down factors begin to affect its behavior (Marslen-Wilson, 1987, p. 78).

On the other hand, once the possibility of top-down information flow is admitted into the system, it is not clear that there are strong empirical grounds for prohibiting it to occur until after the current stimulus is presented. To the contrary, there are several reasons to think that predictive top-down mechanisms would be useful for a system to have: such mechanisms could allow activation processes to reach threshold faster; they could allow the system to maximize processing capacity by constructing representations ahead of time, and in cases where parts of the input stream are missing, such mechanisms would allow the system to construct a good guess of what the input was without any external input. Therefore, at several points in the text I will make the assumption that evidence for top-down effects in sentence processing is support for predictive mechanisms. As I discuss in more detail in Chapter 7, it has proved difficult to acquire the timing information necessary to prove that a top-down effect is predictive. However, in the final chapter I will describe preliminary results from other groups that suggest that the system at least has the capacity for prediction, although it remains to be shown that predictive mechanisms are responsible for particular cases.

1.4 Where do linguistic predictions come from?

Inherent to the idea of a system that makes predictions is that the system has knowledge of contingencies that hold between inputs or, correspondingly, between the abstract representations that the inputs are associated with. In some cases it is easier to determine how the contingencies themselves are represented than in others. For example, in simple learned stimulus-association cases, one could imagine that the probability of a stimulus being followed by a particular target in that task is simply stored and used to generate predictions. Similarly, lexical influences on phoneme recognition (e.g. the increased tendency to hear an ambiguous sound as a /p/ if it is followed by 'ennsylvania') have been modeled as top-down effects mediated through stored connections between the words and the phonemes that compose them (McClelland & Elman, 1986). However, most lexical predictions based on sentence context of the kind that I will talk about in much of the dissertation cannot be captured by stored direct connections between the entire linguistic context and the target because even sentence contexts that have never been heard before can create strong predictions, and such predictions also cannot be captured by direct connections between aspects of the linguistic context and the predicted target, because they arise through interactions between parts of the context. Take the example sentence context *It was a windy day, so the boy went out to fly his…* This context strongly predicts that the next word will be *kite*, as measured by sentence completion norms (DeLong et al., 2005). One could try to argue that this prediction is due to a simple lexical-lexical association between *windy* and *kite*, but this prediction seems to be absent in other sentences that contain *windy*, such as *It was a windy day, so the boy went to the…* It seems to be the case that it is an interaction between a number of elements of the context that gives rise to the prediction, based in some kind of knowledge about the world and knowledge about what kinds of propositions are likely to be expressed in language, but because the representation of conceptual and world knowledge is such a difficult problem it is harder to imagine how such contingencies are represented and computed.

In this dissertation I am going to take it for granted that we can represent such complex contingencies, as is supported by data showing that for many sentence contexts similar to the *kite* example, participants share the same intuitions about which completions are most likely. Here I will not try to solve the interesting but difficult problem of how knowledge of these contingencies are represented; I will focus instead on how knowledge of these contingencies impacts processing of bottom-up input.

1.5 Outline of the dissertation

Chapters 2, 3, and 4 focus on top-down mechanisms underlying access of stored lexical information, while Chapters 5 and 6 focus on top-down mechanisms involved in syntactic processing.

Chapters 2 and 3 examine whether a well-known context effect—differences in the amplitude of the N400 ERP component—can be interpreted as an index of top-down or predictive facilitation, rather than as an index of integration difficulty as has often been assumed. In Chapter 2, I present an MEG study that directly compares N400 effects in semantic priming and sentence context manipulations. I find that the neural response to a word stimulus in the 250-500 ms window varies with the degree to which the context predicts that word, whether or not that context supports construction of larger derived representations. This pattern suggests that the differences in the neural response reflect contextual influences on the mechanisms involved in accessing and/or selecting the word itself, and not the mechanisms involved in entering that word into larger derived representations or assessing the well-formedness of the derived representation.

In Chapter 3, I follow the example of research on vision by using localization as a tool to determine whether an effect of context involves top-down information flow. I present results from a combined fMRI-MEG study and from a meta-analysis of localization studies examining contextual effects in language comprehension. I find that the degree to which the context predicts the stimulus is associated with differential activity in left posterior middle temporal cortex across many functional neuroimaging studies. There is independent evidence that this region supports long-term memory storage of lexical information. Therefore, this result also suggests that context influences the mechanisms involved in accessing the stimulus itself, and that top-down or predictive mechanisms change the state of long-term representations of lexical information.

In Chapter 4, I use evidence from the meta-analysis as a starting point for a neuroanatomical model designed to account for contextual effects in language comprehension. I focus in particular on the role of left inferior frontal cortex in implementing top-down effects on the representational level. The meta-analysis presented in Chapter 3 showed that differential activity in inferior frontal cortex is associated with variation in contextual support only when there is greater than at least 200 ms between presentation of context and target. Following previous authors, I review many studies showing that differential activity in anterior inferior frontal cortex is associated with situations in which specific semantic information about a word or object must be retrieved from memory, and that differential activity in posterior inferior frontal cortex is associated with situations in which one among several activated representations is selected based on context or task. Based on these results, I hypothesize that anterior inferior frontal cortex supports pre-activation of long-term memory representations of lexical and conceptual information, and that posterior inferior frontal cortex supports selection when the pre-activated representations and the representations supported by bottom-up information conflict, while separate areas support construction and maintenance of derived representations.

In Chapter 5, I present an ERP study on syntactic category prediction. I find that the neural response to a function word that constitutes a phrase structure violation (cannot be added to the current phrase structure within the constraints of the grammar) is greater when the syntactic context requires a word of a particular category to appear in the future. Although several interpretations of this result are possible, it can be taken as preliminary support for the hypothesis that phrase structure requirements are instantiated as syntactic predictions.

In Chapter 6, I turn to another way in which prediction may facilitate the process of parsing syntactic dependencies. Across a number of studies with colleagues, we have found that processing of syntactic dependencies is more accurate in syntactic contexts where the first element of the dependency allows prediction of the second element of the dependency. I suggest that this difference in accuracy can be explained if the prediction of the second element of the dependency allows a simple prediction-target match process to take the place of error-prone memory retrieval processes when the second element of the dependency is encountered.

Finally, in Chapter 7, I discuss several issues pertaining to testing and modeling predictive mechanisms more generally, and review general conclusions.

2 The N400 effect as an index of lexical prediction - I

2.1 Introduction

In the next three chapters I will present evidence that top-down mechanisms support the access of stored lexical-semantic representations. This is by no means a new endeavor; the role of contextual information in lexical access has been a major concern of language processing research over the past several decades. Results from many behavioral studies showing contextual effects on phonemic and/or lexical tasks (Ladefoged & Broadbent, 1957; Warren, 1970; Fischler & Bloom, 1979; Ganong, 1980; Stanovich & West, 1981; Elman & McClelland, 1988; Duffy et al., 1989) can be taken as evidence that top-down information based on the context influences the access process. However, as I will discuss below, such results can also be interpreted as post-access effects, in which the contextual information impacts only integration, decision, or response processes (e.g. Norris, McQueen, & Cutler, 2000). Certain measures have been argued to be less affected by response processes than others (i.e., naming vs. lexical decision), but it has been difficult to identify a measure that solely reflects effects on the access mechanism.

Electrophysiological measures, which do not require a behavioral response, provide a means of studying the role of contextual information on access without the interference of overt decision or response processes. Indeed, the ERP response to words known as the N400 component has been shown to be consistently modulated by both lexical and contextual variables, and would thus seem to provide an excellent tool for examining the time course of top-down information in lexical access. However, this electrophysiological measure currently retains the same problems of interpretation as the behavioral literature: while on one view, differences in N400 amplitude reflect ease of lexical access due to priming or pre-activation (Kutas & Federmeier, 2000; Federmeier, 2007), on another the N400 effect reflects the relative ease or difficulty of integration (Osterhout & Holcomb, 1992; Brown & Hagoort, 1993; Hagoort, 2008). Without consensus on the functional interpretation of the N400 effect, the component cannot be used unambiguously as a means of answering questions about mechanisms of access or integration.

This is the first of two chapters that will argue that the N400 effect reflects pre-activation of lexical representations and thus can be used as a tool in future investigations to better understand predictive mechanisms in language. In this chapter, I provide evidence from a within-subjects MEG comparison of two paradigms that share a potential for contextual prediction but that differ in the degree to which integration is required. In the next chapter, I use evidence about where the N400 is generated to shed insight on the functional interpretation of the effect.

2.2 The difficulty of identifying top-down effects

Finding evidence for top-down mechanisms is difficult for several reasons. First, contextual support is confounded with context-target congruity in many designs. For example, a classic finding is that lexical decisions and naming are faster to words that follow a congruous sentence context (1) than to words that follow an incongruous sentence context (2) (Schuberth & Eimas, 1977; West & Stanovich, 1978).

(1) The skier was buried in the *snow*.
(2) The bodyguard drove the *snow*.

A top-down account of these effects would be that the contextual support for the target in (1) facilitated access of the target, speeding reaction times. However, it could equally be the case that the reaction time effect is entirely carried by a slowdown in (2), due to a clash between context and input that is computed after the target is accessed and integrated with the context.

One way to deal with this concern is to compare the congruous-predictive and incongruous-unpredictive contexts against a 'neutral', or congruous-unpredictive, baseline. For example, (Stanovich & West, 1983) used a neutral frame context, as in (3).

(3) They said it was the *snow*.

In theory, if contextual effects are due to predictive pre-activation, processing should be faster in (1) than (3), while if contextual effects are due to context-input mismatch difficulty, processing should be slower in (2) than (3). However, the choice of a truly neutral context raises its own difficulties, as neutral frames are often intuitively less attention-grabbing, and if repeated throughout the experiment as is often the case, may lead to additional differences in attention and strategic processing.

A second concern in interpreting contextual effects as top-down effects is that the higher-level context may impact activity relating to the response, but may not actually impact activity at the representational level in question. For example, a number of studies have demonstrated that the lexical environment in which a phoneme appears affects its identification; if a sound ambiguous between /p/ and /b/ is presented in the context of _*eel*, it is more likely to be reported as a /p/, because *peel* is a common word but */beel/* is not (e.g., Ganong, 1980). The top-down interpretation of this effect is that when the lexical representation of *peel* is activated, it correspondingly biases activity to the phonemes within that word. However, Cutler, Norris, and colleagues have argued that these results can be explained without resort to top-down mechanisms (Cutler & Norris, 1979; Norris, McQueen, & Cutler, 2000). Cutler and Norris (1979) suggest that there are two levels of phonemic representation, one that represents the input and one that represents the system's 'guess' based on both the input and the context. Norris, McQueen, and Cutler (2000) suggest that there is a level of representation essentially dedicated to phoneme identification, and that it is activity at this level and not at the phoneme level that receives input from the context. The force of these arguments depends on the assumption that the task allows use of a distinct representational level dedicated to the response. This might make sense for somewhat unnatural tasks like phoneme identification, but to the extent that top-down effects are task-independent, particularly in cases where no immediate response is required as in many electrophysiological designs, such an interpretation seems less plausible.

The current experiment tries to eliminate these two difficulties by a) comparing sentence contexts that confound contextual support and congruity with single-word contexts for which congruity is less obviously a concern; and b) using an electrophysiological measure that obviates the need for an immediate response on the target. Before introducing the experiment, I pause in the next section to review the key properties of the N400 component in ERP and the manipulations that modulate its amplitude. Moreover, I address several widely-held misconceptions: 1) this large negative deflection in the ERP must reflect a single process; 2) effects on the N400 occur at around 400 ms, too late to reflect anything but strategic processes; 3) differences in

N400 amplitude necessarily reflect increased amplitude in more difficult conditions; 4) differences in N400 amplitude are mainly driven by semantic anomaly.

2.3 The N400 component

During the 1980s, it was observed that after presentation of a number of different kinds of stimuli—words, faces, pictures—a characteristic pattern in the ERP (averaged over many trials) could be observed: from around ~250 ms, average activity began to increase in negativity, until around 400 ms, when activity began to increase in positivity again until around ~500 ms (see e.g., Kutas, Van Petten, & Kluender, 2006 for review). This pattern is what we now refer to as the 'N400 component' (Figure 1).

A couple of points are important to emphasize about the N400 component because our intuitions about waves are often misleading. These intuitions have had important consequences for how researchers have thought about the N400 contextual effect.

First of all, although the N400 component is characterized by 'increasing negativity' in the ~250-400 ms time-window, the actual sign of the activity in the ERP is not always negative, as Figure 1 illustrates. Because there are various methods of measuring voltages at particular electrodes[2], the distance of the ERP from zero is only meaningful in that the ERP waveforms presented have usually been baselined, so that the distance from zero reflects the distance from the activity at that site in an interval before the stimulus was presented. Because the amount of activity in this baseline period

---

[2] E.g., with reference to activity recorded at specified 'reference' electrodes placed on the nose or mastoid, or with reference to some summary measure of the activity across all recording sites

depends on presentation parameters and task requirements, the absolute value of the ERP can differ widely across experiments.

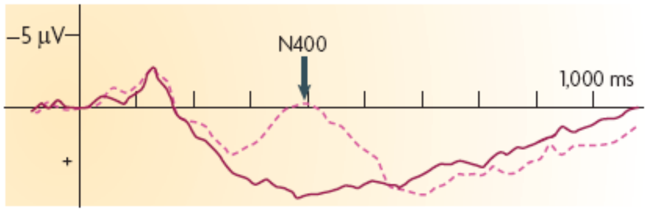


Figure 1. The standard N400 effect in sentential context. The figure illustrates the average event-related potential (ERP) (n=10) recorded from a central midline electrode following the onset of the critical final word in visually presented sentences in which the final word is strongly predicted by the rest of the sentence, as in *I like my coffee with cream and sugar.* The strength of this prediction is determined by a 'Cloze' procedure in which participants are asked to provide endings to sentence contexts. The solid line shows the response to predicted endings, such as *sugar*, and the dotted line shows the response to unpredicted, semantically incongruent endings, such as *socks*. The arrow indicates the N400 component, which is clearly more negative for the incongruent ending.

Second, the local voltage peak that is observed at ~400 ms does not indicate a peak in *brain* activity at 400 ms. More generally, *no* salient visual features of a raw ERP can be given a meaningful interpretation without independent evidence from direct comparisons of ERPs between conditions (Luck, 2005). This is very obvious from a mathematical standpoint but goes strongly against many people's intuition, so it is worth thinking about carefully. The reason peaks in raw ERPs are not meaningful is because we know that the ERP at each particular electrode represents the sum of activity across a number of brain areas, and a peak could be the result of summing any number of different waves (see Figure 2). For example, the N400 wave could be the sum of three sources of activity: one process beginning at 100 ms and dropping off around 300 ms that generates activity that is positive with respect to the reference; one process beginning at 200 ms and continuing at the same level until about 800 ms that generates activity that is negative with respect to the reference; and one process beginning at 400 ms and continuing until about 1000 ms that generates activity that is positive with respect to the reference. Note that in this world, none of the processes are at their maximum at 400 ms, but they create a waveform that is. That is a problem if the waveform is interpreted as a straightforward correlate of a single cognitive process; as for example when it is argued that activity reflected by the N400 component is too late to reflect lexical processing because it happens ‘at’ 400 ms (e.g. Sereno & Rayner, 2003).

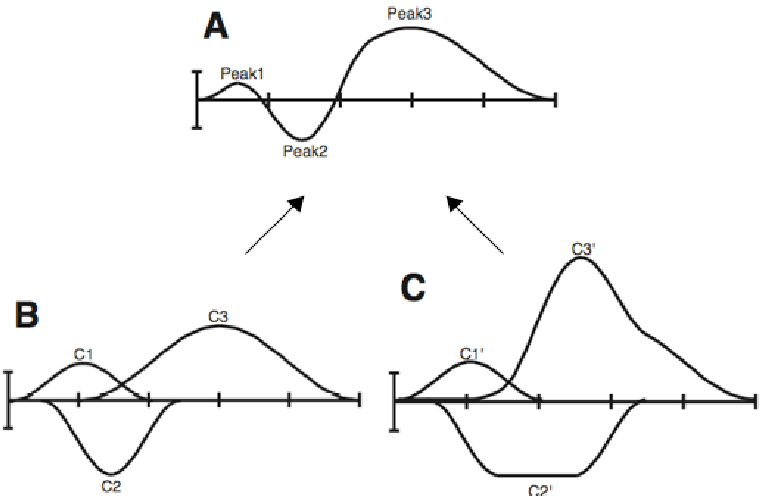


Figure 2. Taken from Luck (2004), this figure illustrates the multiplicity of possible underlying components for an ERP. (A) depicts an ERP to some stimulus, while (B) and (C) depict two (of many) possible arrangements of underlying components which would sum to give the observed ERP.

The design of most ERP experiments avoids this problem in interpretation by comparing the difference between two different experimental conditions. Peaks in ERP difference waves are meaningful, particularly if there are good theoretical reasons to think that the response to the two different conditions differs in only one component process; a peak would then at least suggest that the brain activity associated with that particular process peaks at that point in time. Even so, the natural tendency to assume that the shape of the raw waveform is straightforwardly interpretable frequently slips back in, perhaps in part because raw waveforms tend to be presented more often than difference waveforms.

For example, it is sometimes reported in reviews—including my own (Lau et al. 2008)—that the N400 component is a response to potentially meaningful stimuli, as it is observed following auditorily or visually presented words, word-like nonwords (Bentin, McCarthy, & Wood, 1985), faces (Barrett & Rugg, 1989), pictures (Barrett & Rugg, 1990; Willems, Ozyurek, & Hagoort, 2008), and meaningful environmental sounds (Van Petten & Rheinfelder, 1995; Orgs, Lange, Dombrowski, & Heil, 2008). However, there are actually relatively few studies that examine stimuli that do not elicit a broad negativity in this general time range, and the results of these studies are inconsistent; although consonant strings were initially argued not to show such a negativity (Holcomb & Neville, 1988; Bentin et al., 1999), several recent studies find evidence for a response to consonant strings similar to words in certain contexts (Laszlo & Federmeier, 2008). Second, the fact that a number of different stimuli show a negative deflection with an onset in the range of 200 ms post-presentation is far from strong evidence that these responses have any shared functional locus. Just to illustrate this, here is an alternative account: early stages of sensory processing take about 150 ms for any stimulus, and afterwards various kinds of processing happens all across the temporal lobe, which is completely different and subserved by completely different areas for different stimuli classes such as visually presented words, auditory presented words, pictures, and faces. The only thing left to explain is why all these different responses are negative and not positive, which could well be due to some boring physical property that would, for example, cause activity across all of the temporal lobe to be negative with respect to the mastoid reference that is most commonly used. On this alternative account, all you could say about the N400 component is that it reflects a lot of temporal lobe processing.

In fact, the truth is probably in between; processing of these stimuli is likely to share some characteristics and not others. In Figure 3, I present the ERP response to the same two conditions (semantically primed word and unprimed word) in visual (left) and auditory (right) modalities. Although there is some similarity in the shape of the waveform between 200 and 500 ms for the unprimed condition across the two modalities, there are also clear differences. Consistent with the idea that the broad negativity reflects a number of subprocesses, the spatial distribution of both the N400 component and contextual manipulations of the response have been shown to differ slightly across different classes of stimuli (e.g. Barrett & Rugg, 1990; Kounios & Holcomb, 1994; Holcomb & McPherson, 1994; Ganis, Kutas, & Sereno, 1996; Federmeier & Kutas, 2001; Van Petten & Rheinfelder, 1995). This suggests either that the N400 component reflects similar computations instantiated in different cortical tissue for different stimulus types (Kutas, Van Petten, & Kluender, 2006) or that the N400 component itself may be composed of a number of sub-responses, some of which are common to different types of meaningful stimuli and some which are not.

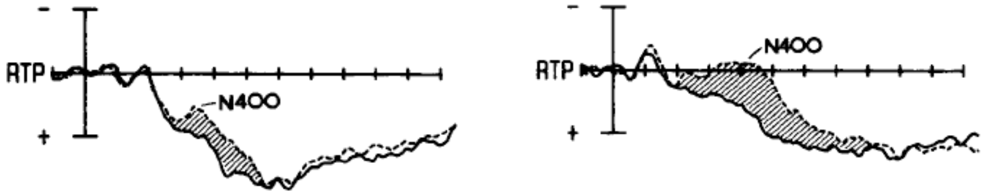


Figure 3. The response to a semantic priming manipulation tested separately in two modalities from Holcomb and Neville (1990). The two figures represent the ERP at an electrode placed over right temporo-parietal areas; each tick on the x-axis represents 100 ms. The dotted line represents the response to a target unrelated to the prime, while the solid line represents the response to a target semantically related to the prime. The left figure presents the results from the visual modality, and the right figure presents the results from the auditory modality.

In this and subsequent chapters, I will be interested in the effect of lexically and semantically predictive context on the ERP response to words. The reason I have spent some time discussing the literature on the N400 is that the most frequently reported type of variation in the N400 is a modulation in amplitude somewhere between ~200-500 ms, in roughly the same time window as the peak of the N400 component in the raw waveform, and therefore is known as the N400 effect in the literature. I will continue to use this terminology here for consistency. However, given the above discussion, it is important to realize that there are probably many functionally distinct 'N400 effects', in that many separate processes contribute to the N400 component and modulation of any of them would lead to modulation of the ERP amplitude during this general time-window. This latter point has not been sufficiently appreciated in the literature, where there have been numerous attempts to identify single causes (and often single neural sources) of different effects during the N400.

To try to ensure that I do not generalize across unrelated 'N400 effects', I will focus on two paradigms that could potentially demonstrate effects of contextual prediction: semantic priming and sentential context. In Section 2.5, I report an MEG experiment that provides evidence that the N400 effect in these two paradigms is due to modulation of the same underlying processes.

2.4 The N400 effect

The contextual modulation in amplitude of the N400 response to words has driven much of ERP language research. As mentioned above, several standard paradigms have been used to show that the N400 response to words is strongly dependent on contextual information. The first and most famous is semantic anomaly: When a sentence is completed with a highly predictable word (*I like my coffee with cream and sugar*), the amplitude of the N400 is much smaller than when the same sentence is completed with a semantically incongruous word (*I like my coffee with cream and socks*; Kutas & Hillyard, 1980). A second frequently used paradigm is semantic priming: When a word target is preceded by a semantically associated word (*salt – pepper*), the amplitude of the N400 is smaller than when a word target is preceded by an unrelated word (*car – pepper*) (Bentin et al., 1985; Rugg, 1985). These effects parallel reaction time data for similar behavioral paradigms in which naming or lexical decision of words in supportive semantic contexts are processed faster than words in unsupportive contexts (Schuberth & Eimas, 1977; West & Stanovich, 1978; Fischler & Bloom, 1979). Throughout the dissertation I will refer to this contextual modulation of N400 amplitude as the *N400 effect*, to distinguish it from the component itself.

What is the functional interpretation of this contextual effect? Many years of research have allowed us to rule out certain possibilities. N400 amplitude seems not to reflect degree of semantic anomaly or implausibility per se, as less expected sentence endings show larger N400 amplitudes than expected ones even when they are both equally plausible (e.g. *I like my coffee with cream and Splenda*; Kutas & Hillyard, 1984). The N400 effect also seems not to be reducible to simple associative priming between words; although lexical association has a reliable effect within sentences (Van Petten, 1993; Van Petten, Weckerly, McIsaac, & Kutas, 1997; Federmeier, Van Petten, Schwarz, & Kutas, 2003) these effects are severely reduced or eliminated when lexical association and congruency with the larger sentence or discourse context are pitted against each other

(Camblin, Gordon, & Swaab, 2007). Furthermore, strong effects of implausibility on sentence endings are observed even when lexical association is controlled (Van Petten, Kutas, Kluender, Mitchiner, & McIsaac, 1991) and when the local context is the same and it is only the larger discourse context that leads to incongruity (St. George, Mannes, & Hoffman, 1994; Van Berkum, Hagoort, & Brown, 1999; Camblin et al., 2007). Finally, N400 amplitude differences seem not to reflect a mismatch response to a violation of the predicted sentence ending, as the amplitude of the N400 response to unexpected sentence endings is not modulated by the strength of the expectation induced by the context (e.g. strong bias context – *The children went outside to* <u>*look*</u> *(*preferred*:play)* vs weak bias context – *Joy was too frightened to* <u>*look*</u> *(*preferred*:speak)*; Kutas & Hillyard, 1984; Federmeier, 2007).

Despite these advances in our understanding of what the N400 effect does *not* reflect, several accounts are still compatible with the existing data. On one view, N400 amplitude is an index of processes of semantic integration of the current word with the semantic and discourse context built up on the basis of previous words (Brown & Hagoort, 1993). On this view, increased N400 amplitudes reflect increased integration difficulty of the critical word with either a prior sentence context or with the prime word. On an alternative view, N400 amplitude indexes processes associated with basic lexical activation and retrieval (Kutas & Federmeier, 2000). On this view, reduced N400 amplitudes reflect facilitated lexical access when the word or sentence context can pre-activate aspects of the representation of the critical word. Hybrid views argue that the N400 actually reflects a summation of several narrower component processes (van den Brink, Brown, & Hagoort, 2001; Pylkkänen & Marantz, 2003). These positions thus parallel those in the older debate in the behavioral literature over whether the locus of context effects on reaction times is pre- or post-lexical in nature; the access view, like the original pre-lexical view, argues that the contextual effect is due to top-down facilitation of access, while the integration view, like the original post-lexical view, argues that the contextual effect is due to processes that occur after the lexical representation has been accessed, in this case integration.

Evidence on the interpretation of the N400 effect has so far been inconclusive. Supporters of the integration view have argued that the N400 effects due to discourse context suggest that the effects reflect the difficulty of integrating the incongruous word into the discourse, but supporters of the access view can account for this by arguing that the discourse predicts or facilitates processing of the congruous ending. Early findings that masked priming paradigms did not show N400 amplitude effects of semantic relatedness (Brown & Hagoort, 1993; Kiefer, 2002; Grossi, 2006) were also taken as evidence that the N400 effect reflected integration with context (and so would not be observed when the context was not consciously perceived). However, later work has suggested that masked priming can affect N400 amplitude under appropriate design parameters (Kiefer, 2002; Grossi, 2006). It has also been argued that, at 400 ms post-stimulus onset, the N400 component is too late to reflect the access process (Hauk, Davis, Pulvermüller, & Marslen-Wilson, 2006; Sereno, Rayner, & Posner, 1998); however, although the context effect peaks at 400 ms, it usually onsets much earlier (~200 ms for visual stimuli), and it is a standard assumption of any view that treats lexical access as an activation process that lexical access is extended over time rather than occurring at a single point.

Supporters of the lexical view have argued that effects of basic lexical parameters on the N400 like frequency and concreteness (Smith & Halgren, 1987; Rugg, 1990; Holcomb & McPherson, 1994), and effects of predictability over and above plausibility (Fischler, Bloom, Childers, Roucos, & Perry, 1983; Kutas & Hillyard, 1984; Federmeier & Kutas, 1999) are more consistent with a lexical basis for N400 effects. However, any factors that make lexical access easier could correspondingly be argued to make semantic integration easier (Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005). Although the semantic priming manipulations that lead to N400 effects could not involve the kinds of syntactic and semantic integration available in sentences, supporters of the integration view argue that the effects are due to less structured integration of prime and target, such as simple matching processes (Brown & Hagoort, 1993). Finally, recent MEG studies using distributed source models provide some support for the view that the context effect reflects multiple processes (Halgren et al., 2002; Maess, Herrmann, Hahne, Nakamura, & Friederici, 2006; Pylkkänen & McElree, 2007), but disparities across studies have made it difficult to assess whether the multiple sources reported contribute to the N400 effect and not other parts of the response, such as the ‘post-N400 positivity’ (Van Petten & Luka, 2006).

In Chapter 3, I will use data on the neural sources of semantic priming and sentential context effects to argue that the N400 contextual effect elicited by these paradigms reflects lexical prediction. However, while the two paradigms have a similar potential for lexical facilitation through contextual pre-activation, they differ in the degree of integrative processes that can be carried out between context and target word. Therefore, it remains possible that the N400 effects that have been observed in the two paradigms reflect different mechanisms that happen to occur in the same time interval, and therefore that their effects must be considered separately.

To assess this possibility, we conducted an MEG experiment, described below, in which we directly compared effects across these context types within participants. To the extent that the timing and topography appears qualitatively similar, we can assume that the same mechanism(s) are driving effects in both paradigms. Furthermore, this design allowed us to test an apparent prediction of the integrative account: since words within sentence contexts presumably require more extensive integration than words in lists or pairs, sentence endings should show a larger N400 than words in pairs, all else being equal. This experiment was conducted in collaboration with Diogo Almeida, Paul Hines, and David Poeppel, and has been previously reported in a submitted manuscript (Lau, Almeida, Hines, & Poeppel, submitted).

2.5 Experiment 1: Word and Sentence Contexts in MEG

A number of studies have compared the paradigms of semantic priming and sentence context using ERPs. Kutas (1993) compared the N400 effect for sentential contexts (expected plausible endings vs unexpected plausible endings) and the N400 effect for single-word contexts (semantically associated vs unassociated words). While the size of the effect was larger for sentential contexts, replicating previous findings, Kutas found that the latency and scalp distribution of the effect was indistinguishable for the sentential and single-word contexts. Nobre & McCarthy (1994) reported subtle differences between the paradigms with a larger electrode array, but as they did not use the same participants across the different context manipulations, this conclusion is somewhat less reliable. These studies were also subject to the concern that the sentence contexts may have confounded both lexical priming and sentential integration effects, even though care was taken to avoid including lexical associates in the sentence materials. Van Petten (1993) mitigated these concerns in a seminal study in which she isolated lexical and sentential context effects by contrasting the effect of lexical association in congruent sentences and in syntactically well-formed nonsense sentences. She found that lexical and sentential context effects thus isolated had a similar scalp distribution and indistinguishable onset latency, although the sentential context effect lasted longer, into the 500-700 ms window.

Although the ERP studies have thus largely supported the hypothesis that lexical and sentential context effects are due to the same underlying mechanism, one might argue that the limited spatial resolution of ERP has caused researchers to miss differences in the neural generators that give rise to these effects. The current study was designed to address this concern. Previous studies have established an MEG correlate to the N400 that shows the same time-course and response properties as the N400 observed in EEG (Helenius et al., 1998; Halgren et al., 2002; Uusvuori et al., 2007). In this study, we compared the N400 effect across these two context types using MEG. MEG measurements are subject to less spatial distortion than ERPs, and thus can provide a better test of whether there exist qualitative differences in the distribution of the effect across these context types.

The access and integration interpretations of the N400 effect also make different predictions about the relative amplitude of the component across the conditions of the sentence and word pair paradigms. An account under which N400 amplitude reflects integration difficulty views the N400 effect as being driven by an *increase* in amplitude for anomalous sentence endings, that are clearly difficult to integrate, while an account under which N400 amplitude reflects lexical access views the N400 effect as driven by a *decrease* in amplitude for predictable sentence endings, where access would be facilitated by contextual support. Therefore, early findings that predictable endings show smaller N400s than congruent but less predictable endings (Kutas & Hillyard, 1984) were taken as support for a facilitated access account, as were findings that N400s to words in congruent sentences are large at the beginning of the sentence and become smaller as the sentence progresses and the next word becomes more predictable (Van Petten & Kutas, 1990, 1991; Van Petten, 1993). These studies further showed that in semantically random sentences, N400 amplitude does not change with word position, even though processing the first word, when the sentence is not yet anomalous, should presumably elicit less integration difficulty than the subsequent words (Van Petten & Kutas, 1991; Van Petten, 1993).

Recently it has been suggested that the integration view can also account for apparent facilitation effects, on the assumption that integration of an item with the context is easier when it can be predicted in advance (Van Berkum et al., 2005; Hagoort, 2008). However, our inclusion of both word pair and sentence stimuli in the same session allows an additional test of the directionality of the context effect, through comparison of the unrelated word pair condition with the anomalous sentence ending condition. Even if the N400 effect for sentence completions is partly driven by facilitation of integration in the congruent condition, integration should be more difficult for the anomalous sentence ending, where the final word must be integrated into a semantic and discourse model that clearly violates world knowledge, than in the unrelated word condition, where there is no need to connect the prime and target in a structured way. Therefore, if the amplitude of the N400 response is shown to be bigger for the anomalous sentence ending than the unrelated word target, it would provide novel support for the integration account.

Participants

21 native English speakers (17 women) participated in the experiment (Mean age: 21, age range: 18-29). All were right-handed (Oldfield, 1971), had normal or corrected to normal vision, and reported no history of hearing problems, language disorders or mental illness. All participants gave written informed consent and were paid to take part in the study, which was approved by the University of Maryland Institutional Review Board.

Design and Materials

The experiment was comprised of two separate tasks, a sentence reading task and a word pair task. Each task included two conditions, *contextually supported* and *contextually unsupported*. In the sentence task, this contrast was achieved by using high cloze probability sentences that ended in either expected or semantically anomalous endings. In the word task, this contrast was achieved by using semantically related or semantically unrelated word pairs.

For the sentence task, 160 sentence frames (4-9 words in length) were chosen for which published norms had demonstrated a cloze probability of greater than 70% for the most likely completion word (Bloom & Fischler, 1980; Kutas, 1993; Lahar, Tun, & Wingfield, 2004). Only sentence frames for which the most likely completion was a noun were included. To form the semantically anomalous versions of these 160 sentences, the same 160 sentence-final words were rearranged among the sentences and the resulting sentences were checked for semantic incongruity. The sentence-final target words had an average log frequency of 1.64 and an average length of 5.1 letters. Examples are presented in (4) and (5). On each of two lists, 80 of the sentence frames appeared with the highly expected ending and 80 of the sentence frames appeared with the semantically anomalous ending. No sentence frame or ending appeared more than once on a given list. The two lists were balanced for surface frequency of the final word.

For the word task, 160 semantically associated word pairs were chosen from existing databases and previous studies that showed semantic priming effects (Nelson, McEvoy, & Schreiber, 2004; Holcomb & Neville, 1990). Both members of each word pair (the prime and the target) were nouns. To form the unrelated word pairs, the primes and targets were re-paired randomly and the resulting pairs were checked for semantic unrelatedness. The target words had an average log frequency of 1.66 and an average length of 4.7 letters, similar to the frequency and length of the target words in the sentence task. On each of two lists, 80 of the primes were followed by a related target and 80 were followed by an unrelated target. No prime or target appeared more than once on a given list. The two lists were balanced for surface frequency of the target word.

Following Kutas (1993), we chose an end-of-trial memory probe for the experimental task in order to match the sentence and word parts of the experiment as closely as possible. In the sentence task, following each sentence participants were presented with a probe word and asked whether the word had appeared in the previous sentence. In the word task, following each word pair participants were presented with a probe letter and asked whether the letter had appeared in the previous words. Word probes were taken from various positions of the sentence, but were always content words.

Letter probes were taken from various positions in both the prime and target words. On each list, the number of yes/no probe responses was balanced within and across conditions.

(4) The pigs wallowed in the mud. (contextually supported)
(5) Bees use the nectar of flowers to make mud. (contextually unsupported)

(6) uncle – aunt (contextually supported)
(7) time – aunt (contextually unsupported)

Procedure

Materials for both tasks were visually presented using DMDX software (K. I. Forster & J. C. Forster, 2003). Sentences were presented one word at a time in the center of the screen using RSVP. Presentation parameters were matched for the sentence and word portions of the experiment as tightly as possible. Except for the first word of sentences and proper names, words were presented only in lower-case. In both tasks words were on the screen for a duration of 300 ms, with 300 ms between words, for a total of 600 ms SOA (stimulus-onset asynchrony). Following the offset of the final word of the sentence or the target word, a 700 ms blank screen was presented before the probe appeared, allowing a 1000 ms epoch from the onset of the critical word before the probe was presented in both tasks. In order to make the probes distinct from the targets, probes were presented in capital letters, followed by a question mark. The probe remained on the screen until a response was made. All words in the experiment were presented in 12-point yellow Courier New font on a black background.

In order to maximize attentiveness across the session, the longer sentence task was presented first, and the faster-paced word pair task was presented second. For the sentence task, participants were instructed that after the sentence was complete, they would be presented with a word and asked to make a button-press response indicating whether the word was present in the previous sentence. For the word task, participants were instructed that after each word pair was presented, they would be presented with a letter and asked to make a button-press response indicating whether the letter was present in either of the two words. Participants were allowed up to 3.5 s to make their response. Both parts of the experiment were preceded by a short practice session, and both parts of the experiment were divided into four blocks, with short breaks in between. Including set-up time, the experimental session took about 1.5 hours.

Recordings and analysis

Participants lay supine in a dimly lit magnetically shielded room (Yokogawa Industries, Tokyo, Japan) and were screened for MEG artifacts due to dental work or metal implants. A localizer scan was performed in order to verify the presence of identifiable MEG responses to 1 kHz and 250 Hz pure tones (M100) and determine adequate head positioning inside the machine.

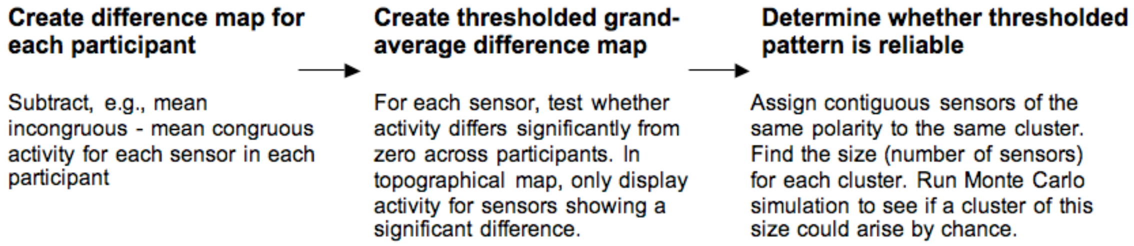
MEG recordings were conducted using a 160-channel axial gradiometer whole-head system (Kanazawa Institute of Technology, Kanazawa, Japan). Data were sampled at 500 Hz and acquired continuously with a 60 Hz notch filter and a 200 Hz high-pass filter. A time-shift PCA filter (de Cheveigné & Simon, 2007) was used to remove external sources of noise artifacts. Epochs with artifacts exceeding 2 pT in amplitude were removed before averaging. Incorrect behavioral responses and trials in which participants failed to respond were also excluded from both behavioral and MEG data analysis. Based on these criteria, 9% of trials were excluded. For the analyses presented below, data were averaged for each condition in each participant and baseline corrected using a 100 ms pre-stimulus interval. For the figures, a low-pass filter of 30 Hz was used for smoothing.

Two participants were excluded from the analysis based on low accuracy (< 80%) in one or both sessions of the experiment, and one participant was excluded from the analysis based on high levels of signal artifact. Data from the 18 remaining participants (9 from each counterbalanced stimuli list) were entered into the MEG analysis.

A measure reported in many MEG experiments is the peak latency and amplitude of the dipolar pattern of interest across different experimental manipulations. Based on early differences between word and sentence conditions (described below), we conducted a peak latency and amplitude analysis on the M170, a component associated with higher-level visual processing (Tarkiainen et al., 1999, 2002). Because of limited sensor coverage over the most occipital areas, most participants showed only one half of the dipolar pattern in each hemisphere. We selected the 5 most active channels in the 140-200 ms window in each hemisphere, across conditions, for each participant, and we report the average peak latency and amplitude of the RMS of these channels.

The response topography generally observed to written words in the time interval associated with the N400 (300-500 ms) has a similar distribution to the response topography previously referred to as the M350 component response to words, which has been posited to reflect lexical access processes (Pylkkänen, Stringfellow, & Marantz, 2002; Pylkkänen & Marantz, 2003). However, we found that not all conditions in the current experiment elicited a response with this topography. While most participants displayed an M350-like pattern in the 300-500 ms window for the word-pair conditions and the sentence-incongruous condition, most displayed a qualitatively different topography to the sentence-congruous condition. Thus, an analysis comparing peak latency and peak amplitude for selected M350 sensors across all 4 conditions would be inappropriate.

As an alternative measure, we created statistically thresholded difference maps for the sentence and word conditions. For each participant, we created two difference maps, one for sentence-incongruous – sentence-congruous conditions, and one for word-unrelated – word-related conditions. We grand-averaged these individual participant maps to create a composite difference map for each task. However, because head position can vary in MEG, participants may contribute unequally to the differential activity observed at individual sensors. Therefore, we created a statistically thresholded difference map for each task by displaying activity only for sensors at which the difference between conditions was significantly different from zero across participants ($p < .01$). To correct for multiple comparisons across the large number of recording sites, we clustered the significant sensors based on spatial proximity and polarity and then conducted a randomization test of the summary test statistic for clusters of sensors that crossed this initial threshold (Maris & Oostenveld, 2007). For each thresholded difference map, we assigned each significant sensor to a cluster. Sensors within 4 cm of each other that demonstrated the same polarity were assigned to the same cluster. A summary test statistic was calculated for each cluster by summing the t-values of all the sensors in that cluster (t-sum). This test statistic (t-sum) was then recomputed in 4000 simulations in which the condition labels for each participant were randomly reassigned. This generated a random sample of the empirical distribution of t-sum under the null hypothesis of no treatment effect (see Figure 4). If the t-sum values from the clusters obtained in our experiment were more extreme than 95% of the t-sum values calculated by the randomization simulations, the cluster was considered statistically significant.



Create difference map for each participant

Subtract, e.g., mean incongruous - mean congruous activity for each sensor in each participant

Create thresholded grand-average difference map

For each sensor, test whether activity differs significantly from zero across participants. In topographical map, only display activity for sensors showing a significant difference.

Determine whether thresholded pattern is reliable

Assign contiguous sensors of the same polarity to the same cluster. Find the size (number of sensors) for each cluster. Run Monte Carlo simulation to see if a cluster of this size could arise by chance.

Figure 4. Flow chart of analysis steps used to create and test the reliability of statistically thresholded topographical maps of differences in activity between conditions.

Finally, to test the simple directionality of contextual effects (facilitative vs. inhibitory) across different conditions, we computed the grand-average RMS over all the sensors in each hemisphere (75 left-hemisphere channels and 75 right-hemisphere channels; 6 midline sensors were excluded for this analysis) for each condition. This provided a relatively insensitive but conservative measure of the most robust experimental effects. 2 × 2 × 2 ANOVAs (hemisphere × task × contextual support) were conducted on the average waveform for each condition over 200 ms windows: 100-300 ms (M170), 300-500 ms (N/M400), 500-700 ms, and 700-900 ms, followed by planned comparisons.

Behavioral results

Mean response times and accuracy for the probe-detection task are presented in Table 1. A 2 × 2 ANOVA with context type (sentence vs. word) and contextual support (supportive vs. unsupportive) as factors showed a significant effect of context type on response times such that the response times were longer overall for the word pair block ($F(1,17) = 32.3$, $p < .01$). There was also a significant interaction between context type and contextual support ($F(1,17) = 14.6$, $p < .01$); pairwise comparisons showed that this interaction was driven by a significant effect of context for sentences (longer RTs for unsupportive contexts; $F(1,17) = 17.1$, $p < .01$) and the absence of a contextual effect for word pairs ($p > .1$). A binomial mixed effects model using context type and contextual support as factors showed a similar pattern for response accuracy: a significant main effect of context type on accuracy ($p < .01$), with lower accuracy for the word pair task, and a main effect of context type ($p < .01$) and an interaction between context type and contextual support ($p < .01$), both apparently driven by a significant effect of context in sentences ($p < .01$) but not for word pairs ($p > .1$). The main effect of context type across RT and accuracy is likely due to inherent differences in the difficulty of the probe-word task used for sentences and the probe-letter task used for words; I will return to this issue in the Discussion. The selective effect of context on response speed and accuracy in the sentence task is likely due to the stronger expectation set up by the sentential context, which probably resulted in greater disruption for the unsupportive case as shown in previous behavioral studies (e.g. Fischler & Bloom, 1979; Stanovich & West, 1981).

	Sentences		*Word Pairs*	
	Supportive context	*Unsupportive context*	*Supportive context*	*Unsupportive context*
RT	714	738	922	906
Accuracy	0.99	0.98	0.95	0.95

Table 1. Reaction times and accuracy for each of the two end-of-trial probe tasks in Experiment 1, by condition.

MEG results

Figure 5 illustrates the grand-average MEG waveform and topography for the context effect (unsupportive – supportive) for each context type. The context effect was much larger in amplitude for sentences than word pairs, consistent with previous literature (Kutas, 1993), but when the magnitude of the word pair effect is scaled (multiplied by 2) to match the magnitude of the sentence effect, the timing and topographical distribution appear quite similar.

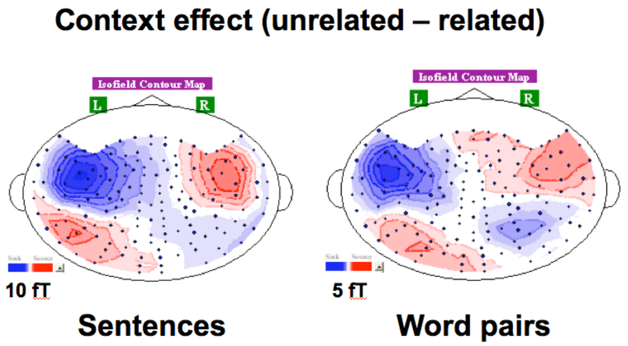


Figure 5. Grand-average whole-head topography of magnetic field potentials for the difference in the response to words following an unsupportive context and words following a supportive context. The activity presented in this image is unthresholded to show general topographical similarities in the contextual effect for the two context types; because the effect was weaker in the word pair conditions, the scale has been decreased for this contrast relative to the sentence contrast.

Comparing effect of contextual support across context types

To compare the contextual effect across context types, it was necessary to select a subset of sensors of interest. As the contextual effect is known to be strongest for sentence contexts, we used the response observed for the sentence conditions as the basis for sensor selection. We created a statistically thresholded topographical map for the unsupportive-supportive sentence contrast (Figure 6). This map displays sensors that showed a reliable difference (*p* < .01) between the two sentence contexts across participants in the 300-500 ms window. This procedure identified three clusters of sensors distinguished by polarity of the effect and hemisphere: a left anterior cluster that showed a negative difference, a left posterior cluster that showed a positive difference, and a right anterior cluster that showed a positive difference. The randomization clustering analysis showed that the three clusters of activity observed in the sentence context map had a less than 5% probability of having arisen by chance (sums of t-values over significant sensors: left anterior sink = -116.1; right anterior source = 54.0; left posterior source = 87.0; 2.5% and 97.5% quantiles = -15.1 > t-sum > 16.0), but that only the left anterior cluster was reliable for the word context effect (sum of t-values = -26.5; 2.5% and 97.5% quantiles = -14.2 > t-sum > 13.6). This could indicate a qualitative difference between context types, but it could also reflect the difference in magnitude between the context types; the word effect may have been too weak to survive the analysis in all but the largest cluster.

In Figure 6 I also present the difference waves for the contextual effect (contextually unsupported – supported) for sentences and word pairs. As the sensors here were defined as those that showed a significant difference between sentence conditions, the fact that the sentence conditions show strong differences in the N400 time-window is unsurprising. The question of interest is whether the word pair conditions also showed a context effect across the same sensors. Indeed, in the word pair comparison we observe a difference between contextual conditions in the 300-450 ms window in all three clusters, significant in the left anterior cluster ($t(17) = 3.66$, $p < .01$), and marginally significant in the right anterior cluster ($t(17) = 2.05$, $p = .057$) and left posterior cluster ($t(17) = 2.04$, $p = .057$). No sensors showed a significant contextual effect for word pairs but not for sentences.

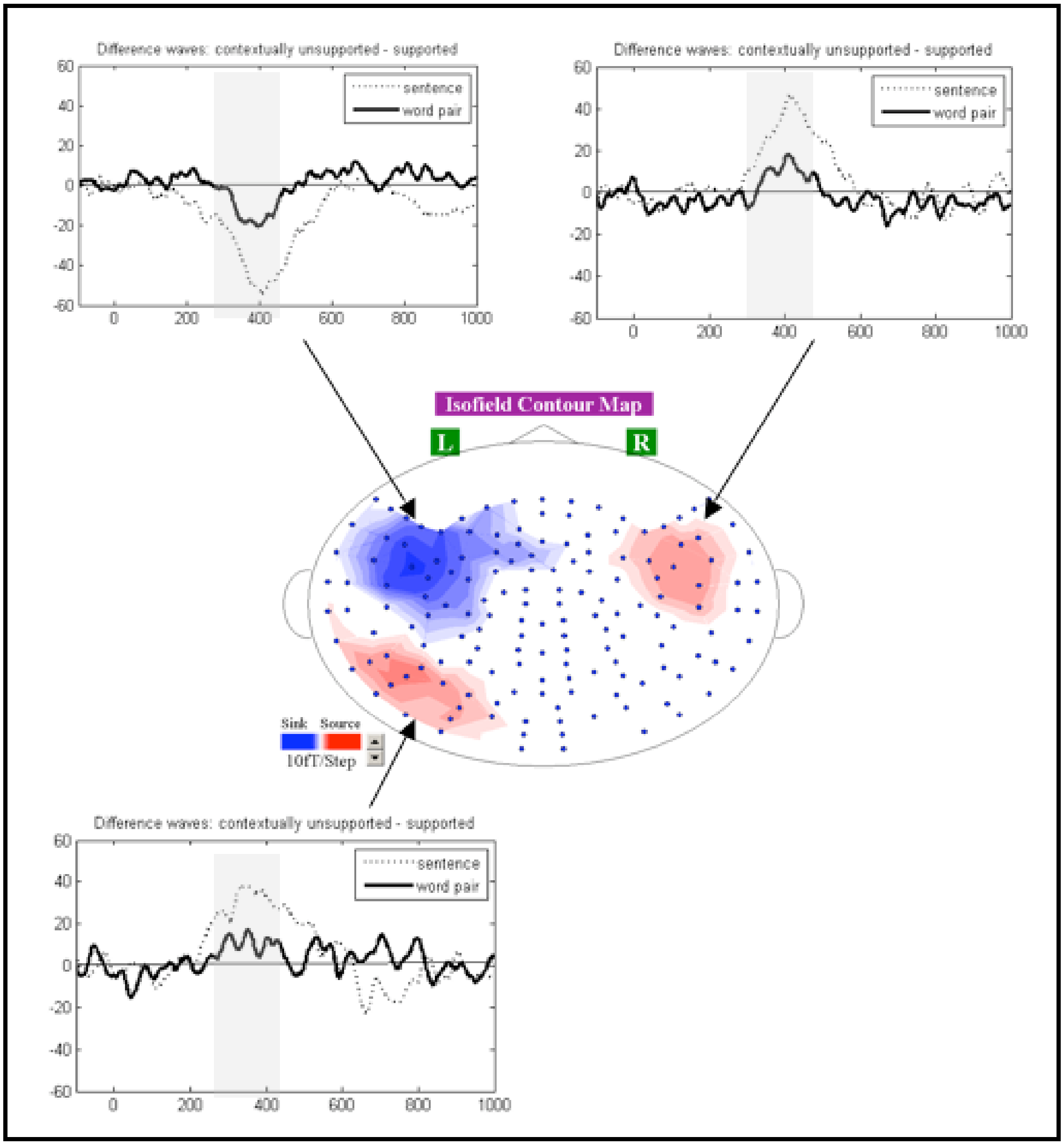


Figure 6. Statistically thresholded grand-average whole-head topography for the sentence ending contrast (contextually unsupported – contextually supported). This image shows only sensors for which the difference was significant across participants between 300-500 ms (*p* < .01). The waveforms show the average difference waves (contextually unsupported – supported) for sentences and word pairs across the significant sensors in the cluster indicated by the arrow; all three clusters were larger than would have been expected by chance (see text).

A further similarity between the sentence and word pair context effects can be observed in the timing of the effects across hemispheres. Visual observation of the grand-average difference map for the sentence condition revealed that the contextual effects observed bilaterally in anterior sensors onset earlier over the left hemisphere than over the right hemisphere (Figure 7). This was confirmed by statistically thresholding the sensors across subjects ($p < .01$); in an early time-window (250-350 ms) only the two left-hemisphere clusters were significant (sums of t-values over significant sensors: left anterior sink = -54.6; right anterior source = 6.6; left posterior source = 83.6; 2.5% and 97.5% quantiles = -16.8 > t-sum > 18.1), while in a later time-window (350-450 ms) the right anterior cluster was also significant (sums of t-values over significant sensors: left anterior sink = -54.6 ; right anterior source = 6.6; left posterior source = 83.6; 2.5% and 97.5% quantiles = -16.9 > t-sum > 17.7).

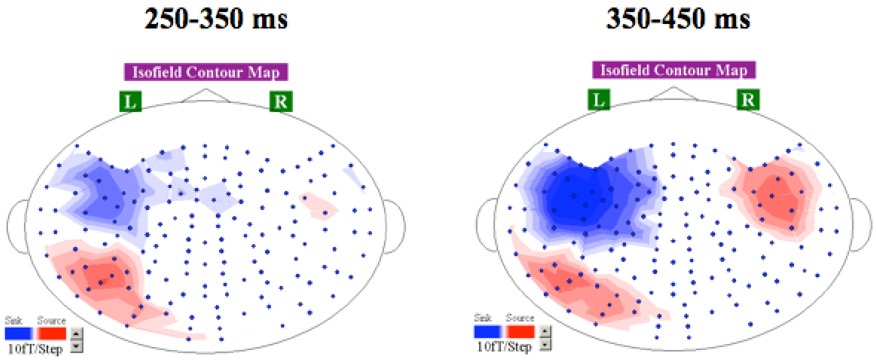


Figure 7. Statistically thresholded grand-average whole-head topography for the sentence ending contrast (contextually unsupported – contextually supported) averaged across two time-windows chosen by visual inspection, showing only those sensors for which the difference between conditions was significant across participants ($p < .01$). In the first time-window, the two left-hemisphere clusters were larger than would have been expected by chance, while in the second time-window, an additional right-hemisphere cluster was also larger than would have been expected by chance (see text).

Crucially, the same right-hemisphere delay was observed in the word conditions. Figure 7 plots the average context difference wave (contextually unsupported-supported) across the left and right anterior clusters depicted in Figure 6, with the polarity of the left cluster waveform reversed to facilitate visual comparison. We tested consecutive 50-ms windows to determine when significant effects of context began ($p < .05$). For both the sentence and the word pair conditions, the contextual effect began earlier in the left anterior sensors than in the right anterior sensors, although effects over both hemispheres became significant later for words than for sentences (200-250 ms (left) vs. 300-350 ms (right) for sentences; 300-350 ms (left) vs. 350-400 (right) ms for words). Importantly, however, both stimulus types showed the same hemispheric asymmetry in latency: left-hemisphere effects of context beginning earlier than right-hemisphere effects.

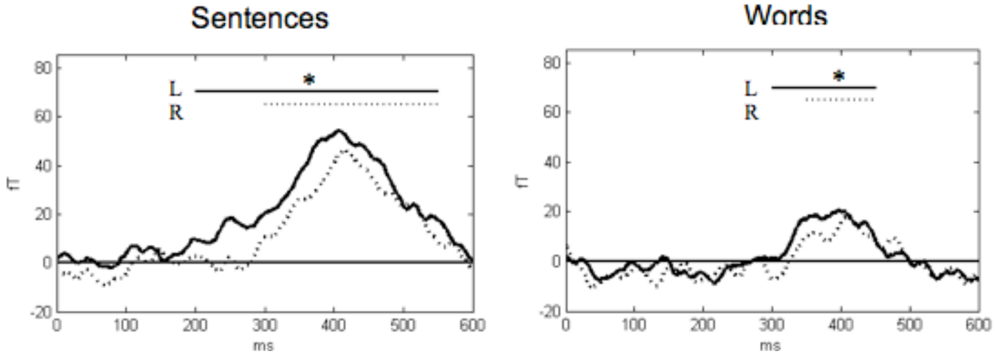


Figure 8. Difference waves (contextually unsupported – supported) averaged across the left and right anterior clusters of sensors depicted in Figure 6 for sentences and words. The polarity of the difference wave in the left cluster has been reversed to facilitate comparison between the timing of the responses.

So far I have presented the waveforms and topographical maps for differences between the contextually unsupported and contextually supported conditions. Figure 9 presents the grand-average MEG waveforms across the two anterior clusters for the individual conditions. This figure indicates that the response to contextually unsupported words was of similar magnitude in word pairs and sentences in left hemisphere clusters, and that the contextually supported words showed a shift towards baseline that was greater in sentences than in word pairs. However, even though the analyses above suggest that the effect of the contextual manipulation is similar in word pairs and sentences, the condition waveforms suggest that the base response to the unsupported condition in the right hemisphere cluster is greater when the target word is embedded in a sentence context. I review possible interpretations of this pattern in the Discussion.

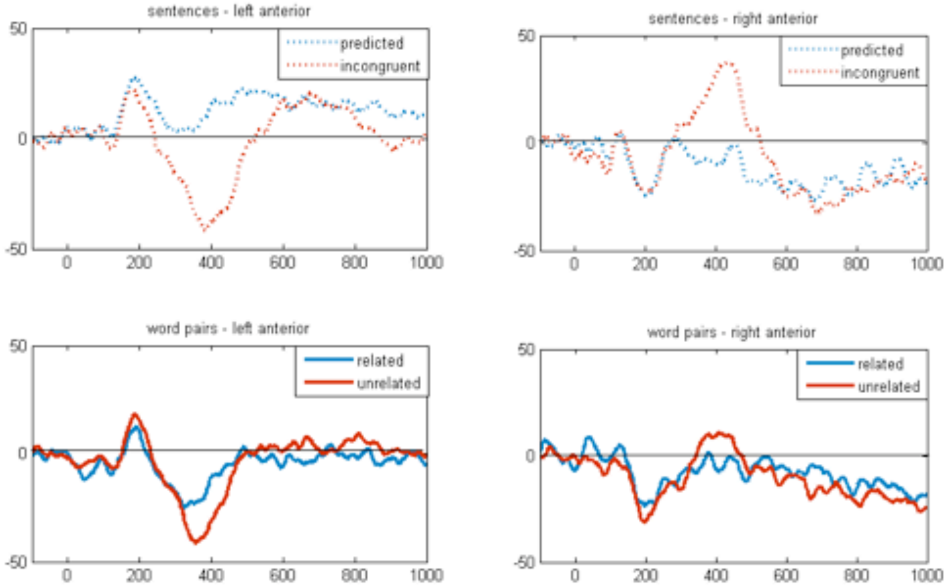


Figure 9. Grand-average MEG waveforms for contextually supported and unsupported words in both sentence (top) and word pair (bottom) contexts, across left and right anterior sensor clusters of sensors that showed a significant effect of context for sentences between 300-500 ms, as defined above.

Overall, consistent with previous ERP studies, we found that sentence and word pairs show similar effects of contextual support in the N400 time window. While the contextual effect for word pairs was much smaller in magnitude, this is plausibly due to differences in the strength of the lexical prediction made possible by sentence contexts and single prime words. The context effect for word pairs was significant across a shorter time window than the context effect for sentences, but this could be due to the smaller magnitude of the effect.

Late effects of context

We also examined activity in a later time window (600-900 ms) within which a post-N400 positivity is often visible in ERP recordings (e.g. Kutas & Hillyard, 1980;

Friederici & Frisch, 2000; Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007; see Van Petten & Luka, 2006 and Kuperberg, 2007 for review), but which has rarely been reported in MEG studies. No significant clusters were found using a threshold of $p < .01$. At a more conservative threshold ($p < .05$) we found two clusters of sensors in the right hemisphere showing a difference between contextually supported and unsupported sentence endings (Figure 10); these clusters were marginally reliable in the nonparametric test (sums of t-values over significant sensors: right anterior sink = -42.7; right posterior source = 32.2; 2.5% and 97.5% quantiles = -34.3 > t-sum > 34.3). No clusters showed a difference between contextually supported and unsupported word pairs.

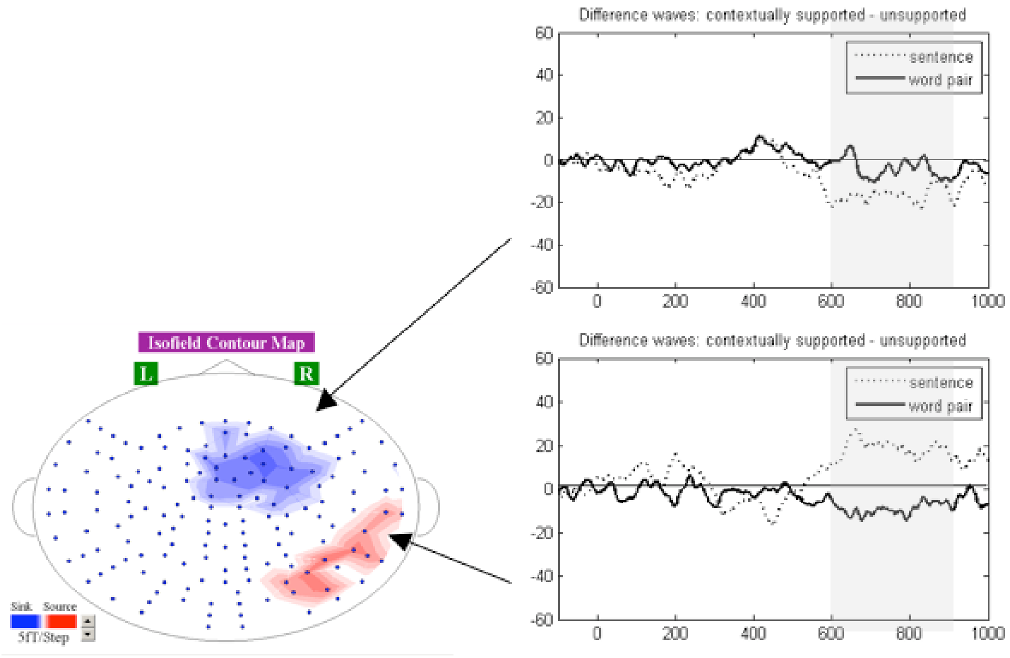


Figure 10. Statistically thresholded grand-average whole-head topography for the difference between contextually unsupported – contextually supported sentence endings. This image only displays activity over those sensors for which this difference was significant across participants between 600-900 ms ($p < .05$). Difference waves for the contextual effect (contextually unsupported – contextually supported) for sentences and word pairs across the sensors in the clusters indicated; these clusters were only marginally reliable, however (see text).

RMS analysis by hemisphere

It is difficult to generalize MEG analyses of subsets of sensors across studies because the position of the head with respect to the sensors is different across participants and MEG systems. Averaging MEG signal strength across all sensors in each hemisphere is one means of avoiding sensor selection and thus making results potentially more generalizable, even though it means that many sensors that certainly do not contribute to the effect of interest will be included, making this a less sensitive measure. Figure 11 presents the grand-average RMS across all sensors over each hemisphere for all four conditions. We observed three patterns of interest, described in more detail below. First, regardless of context, words in sentences showed a larger M170 response than isolated words in right hemisphere sensors. Second, in left hemisphere sensors, the response to incongruous sentence endings was relatively similar in amplitude to both related and unrelated word targets during the N400 time window, whereas the response to congruous sentence endings was strongly reduced. Third, during the same time window in right hemisphere sensors, the response to incongruous sentence endings showed increased amplitude relative to the other three conditions.

In the M170 window (100-300 ms) we found a main effect of hemisphere ($F(1,17) = 11.35$; $p < .01$) and a marginally significant interaction between task and hemisphere ($F(1,17) = 3.47$; $p < .08$). These effects seemed to be driven by a hemispheric asymmetry in M170 amplitude for the word conditions (increased amplitude in the left hemisphere) but not the sentence conditions. Such a leftward asymmetry in the word conditions may have been due to the ‘local’ attention to the letters of the word required to perform the probe task (e.g. Robertson & Lamb, 1991). No other early effects were observed.

In the N400 window (300-500 ms) there were main effects of hemisphere ($F(1,17) = 27.17$; $p < .01$), and contextual support ($F(1,17) = 11.93$; $p < .01$), such that signal strength was greater in the left hemisphere than the right and was greater in the unsupportive context than in the supportive context. There were marginally significant interactions between contextual support and hemisphere ($F(1,17) = 4.05$; $p < .06$) and between context type and hemisphere ($F(1,17) = 3.88$; $p < .07$), and a significant interaction between contextual support and context type ($F(1,17) = 16.03$; $p < .01$). These interactions seemed to be driven by the strong reduction in activity observed in the left hemisphere for the congruent sentence ending relative to the other three conditions and the strong increase in activity observed for the incongruent sentence ending in the right hemisphere (Figure 11). Paired comparisons between the sentence endings and the unprimed word target confirmed this visual impression. In the left hemisphere there was a significant difference between the unprimed word target and the congruent sentence ending ($F(1,17) = 14.0$; $p < .01$) but not the incongruent sentence ending ($F(1,17) = 1.07$; $p > .1$). In the right hemisphere there was no difference between the unprimed word and the congruent ending ($F(1,17) = .02$; $p > .1$), although there was a marginally significant difference between the unprimed word and the incongruent ending ($F(1,17) = 4.11$; $p < .06$), which I return to in the Discussion.

No significant main effects or interactions were observed for the late (600-900 ms) window in which post-N400 positivities are sometimes found (*p*s > .1).

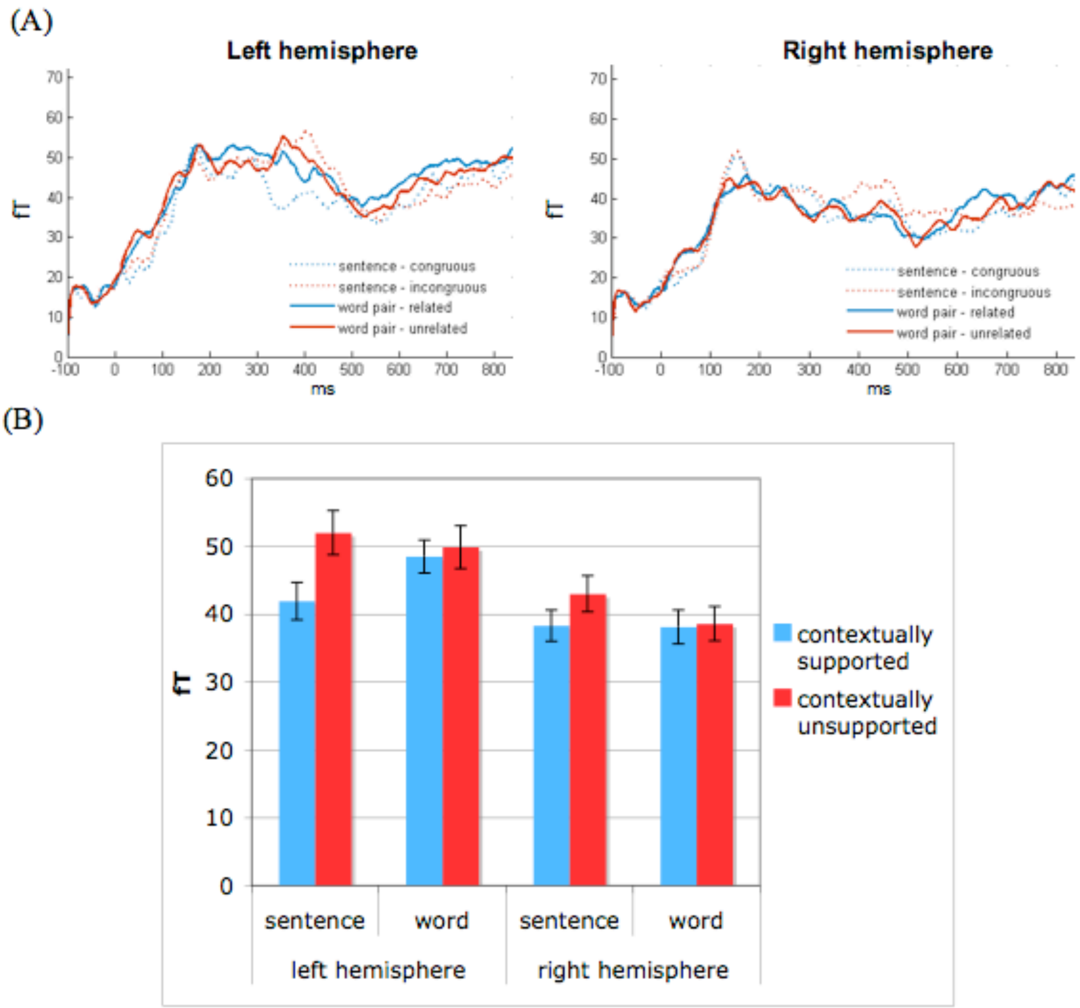


Figure 11. (A) Grand-average RMS across all sensors in each hemisphere (75 sensors in left; 75 sensors in right) for all 4 conditions. (B) Grand-average RMS amplitude across the 300-500 ms window in each hemisphere.

Discussion

In this experiment we used MEG to compare the electrophysiological effects of contextual support in more structured sentence contexts with less structured word pairs in the time window associated with the N400 effect. The aim was to try to separate the contributions of lexical pre-activation (possible in both context types) from integration (possible to a much greater degree in the sentence context). We did find some evidence of two different topographic patterns over the N400 time-window, supporting the idea of multiple contributors to the N400 contextual effect. However, both sentence and word pair contexts showed this same two-phase pattern, although sentence contexts demonstrated contextual effects of greater magnitude.

The finding that the N400 contextual effect in MEG is qualitatively similar for sentence contexts and word contexts is consistent with earlier work by Kutas (1993) that showed that the timing of the N400 effect in EEG and its topographical distribution over the scalp were indistinguishable for the two context types. MEG signals are not subject to the same field distortions that affect EEG, and thus, using MEG provides an opportunity of spatially separating components that could appear to be the same in EEG. Although future studies could conduct more detailed spatial analyses on this type of MEG data to test for even subtler distinctions in the response, the fact that the MEG profile is so similar for N400 effects across the two context types provides compelling evidence that the effects observed across these quite different paradigms do indeed share a common locus.

A second revealing finding was that the N400 response for incongruous sentence endings did not differ from the response to unprimed words over the left hemisphere; rather, it was the congruous or predicted sentence endings that showed a significant difference from unprimed words, in the form of a reduction in activity. If the amplitude of the N400 reflected integration difficulty, one would expect that N400 amplitude should be greater when a word must be integrated with an obviously incongruous highly-structured prior context (*I like my coffee with cream and socks*) than when a word must be integrated with a non-structured, and thus more neutral, prime-word context (*priest – socks*). In the first case, the structure highly constrains the possible interpretations of the sentence, such that any licit interpretation is at least unusual, while in the second case, the absence of structure means that 1) a tight integration of the two words is not required and 2) a number of relations are possible between the two words, increasing the likelihood that a congruous integration of the two concepts can be easily achieved (e.g., *priests wear socks*, *the socks belonging to the priest*). However, the results here show no difference between incongruous sentence endings and unprimed words in left hemisphere sensors, where the N400 effect was earliest and more spatially widespread. The congruous endings, on the other hand, showed a significant reduction in amplitude relative to unprimed words, consistent with lexical facilitation accounts in which the lexical entry for the word forming the sentence ending can be pre-activated by the highly predictive sentence frames used in most N400 studies. This is consistent with previous work suggesting that the N400 effect is driven by a reduction in activity relative to a neutral baseline (Van Petten & Kutas, 1990; 1991; Van Petten, 1993; Kutas & Federmeier, 2000). One caveat to the results presented here, however, is that in order to simplify creation of the materials, different target words were used in the sentence task and word task. Although the target words were similar in frequency and length across tasks, some other difference between the items could have contributed to differences in response. Future work could resolve this concern with a design in which the same words were used as targets in the two different tasks across participants.

Previous authors have suggested that predictability effects not due to congruity can still be explained by an integration account of the N400, if prediction of an upcoming word can facilitate integration processes (Van Berkum et al., 2005; Hagoort, 2008). In other words, prediction can take the form of not only pre-activation of a stored lexical entry, but also pre-integration of the predicted lexical item with the current context. Therefore, if the N400 reflects integration difficulty, it should be reduced in predictive cases where little integration work is left to be done by the time the bottom-up input is encountered. While this hypothesis can explain the reduction observed in the congruent sentence endings relative to a neutral baseline, the lack of an increase in the incongruent sentence endings still seems to be unaccounted for. If the N400 does not reflect the difference between a case where there are few pragmatic constraints on integration and a case where there are strong pragmatic constraints that make a felicitous derived representation hard to achieve, it is hard to see how its amplitude could be said to reflect integration difficulty. At the least, this account requires a re-defining and sharpening of what is meant by ‘integration’ (Van Berkum, in press).

In right hemisphere sensors, we did observe evidence of the pattern predicted by the integration account: a marginally significant difference between incongruous sentence endings and unprimed words, and no significant difference between congruous endings and unprimed words. One potential explanation of this asymmetry is that left hemisphere sensors reflect predictability in the N400 time window, while right hemisphere sensors reflect integration difficulty; Federmeier and colleagues have previously suggested that the left hemisphere may be preferentially dedicated to prediction in comprehension (e.g. Federmeier & Kutas, 2003b; Federmeier, 2007). However, since the size and timing of the context effect was so similar across corresponding left and right anterior sensor clusters (Figure 8), this account would need to assume that predictability yields separate effects on access and integration that are virtually identical in timecourse.

An alternative is that it is not differences in the size of the context effect that drive the hemispheric asymmetry, but differences in the base level of activity for sentences and words. For sentences, sensors in both hemispheres show a strong peak of activity for incongruent endings and activity close to baseline for predicted endings. For words in word pairs, on the other hand, both unrelated and related targets show a broad peak of activity in left hemisphere sensors, but in right hemisphere sensors, activity is close to baseline for both conditions. An explanation for the asymmetry consistent with this pattern is that sources reflected in right hemisphere sensors are recruited to a greater degree during normal sentence processing than in processing of isolated words. If this were true, activity in right hemisphere sensors would *always* be higher in amplitude for words in sentences than words in pairs, all other things being equal. Future work will be needed to determine which of these accounts best explains the asymmetry observed.

Finally, we also observed a number of right hemisphere sensors that showed a significant effect of context for sentence endings in the later part of the processing timecourse, between 600-900 ms, although this effect was less reliable than the earlier effects. The timing of this effect is consistent with a late positivity often observed following N400 contextual effects in ERP that has sometimes been called the *post-N400 positivity* (Van Petten & Luka, 2006). As this effect was observed for the sentence contexts, which differed in predictability and semantic congruity, but not for the word pairs, for which semantic congruity is less well-defined, this later effect may reflect difficulty in compositional semantic integration, reanalysis, or some other response specific to semantic incongruity. However, it could also reflect a sentence-specific mechanism engaged during normal processing that is simply facilitated in predictive contexts. More work must be done to tease these possibilities apart.

2.6 Conclusion

In this chapter I used MEG to show that (1) the effect of contextual support during the N400 time-window is qualitatively similar in timing and topography for words presented in structured sentences and words presented in unstructured word pairs, (2) this time window is associated with at least two distinct MEG topographies, and (3) in left hemisphere sensors, at least, the effect is driven by a reduction in amplitude in the predicted sentence ending relative to an unprimed target word, rather than an increase in amplitude in the incongruous ending. These results suggest that the N400 effect in sentence contexts and semantic priming reflect the same underlying mechanisms, which provides support for combining data from these two paradigms in the meta-analysis presented in Chapter 3. More importantly, the results provide support for the view that at least part of the N400 context effect observed in ERP reflects facilitated access of stored information rather than relative difficulty of semantic integration, although the results also suggest that the effect may reflect more than one mechanism. In the next chapter, I will turn to data on localization to provide a more direct argument that the N400 effect can be interpreted as an index of true top-down effects on the access level.

3 The N400 effect as an index of lexical prediction - II

3.1 Introduction

In Chapter 2 I found that the neural response to a word stimulus in the 250-500 ms window varies with the degree to which the context predicts that stimulus, whether or not that context supports construction of larger derived representations. This pattern suggests that the differences in the neural response reflect contextual influences on the mechanisms involved in accessing and/or selecting the stimulus itself, and not the mechanisms involved in entering that stimulus into larger derived representations or assessing the well-formedness of the derived representation.

The literature on visual processing suggests an interesting complementary approach to disambiguating the source of contextual effects, which is to identify the anatomical location of the representational level in question and determine whether context impacts activity in this area. This is the approach I will pursue in this chapter.

In vision, the cortical substrate for the first stages of visual analysis is fairly uncontroversial. It is known that primary visual cortex (V1 – Brodmann's area 17) is the first cortical area to receive visual information (~40-60 ms post-stimulus onset; Bullier & Nowak, 1995), and that it codes simple visual features like line orientation over tiny parts of the receptive field. After early processing in V1, information is passed to other visual areas such as V2, MT, V4, and inferotemporal cortex, which are associated with representation of more complex visual features such as motion, color, shapes, and objects.

Given this architecture, effects of the larger visual context on activity in V1 cells would implicate recurrent feedback from higher-level cortical areas. A number of studies have therefore localized contextual effects in V1 as a means of providing evidence for top-down information flow. For example, visual responses in V1 cells have been shown to be sensitive to illusory contour stimuli that are only perceived as present by virtue of their context (Lee and Nguyen, 2001). Similarly, Lamme and colleagues have shown that in images containing figure-ground relationships, V1 cells give a larger response to a line of their preferred orientation when it occurs inside of the perceived figure than when it occurs inside the perceived ground, whether the figure-ground relationship is defined by motion, texture, disparity, color, or luminance (Lamme, 1995; Zipser, Lamme, & Schiller, 1996; Lee, Mumford, Romero, & Lamme, 1998). In humans, fMRI studies find that V1 activity is affected by the degree of higher-order complexity in the stimuli, even when the low-level stimulus properties are tightly controlled, which has been argued to be due to predictive feedback from higher-level areas (Murray, Kersten, Olshausen, Schrater, & Woods, 2002; Paradis et al., 2000). The fact that these effects seem to generalize across different types of context and often onset later than the initial response to the bottom-up stimulus argues against the alternative explanation that V1 cells are inherently sensitive to higher-level properties in the absence of feedback (Rossi, Desimone, & Ungerleider, 2001).

In this chapter, I will pursue an analogous approach for disambiguating contextual effects in language processing. If the broader context affects activity in the cortical areas that subserve storage of lexical representations, it would constitute fairly strong evidence that context influences lexical access. Therefore, in this chapter I will consider evidence on the localization of contextual effects on lexical processing. First I will review previous evidence from imaging and aphasia suggesting that mid-posterior middle temporal gyrus (MTG) supports storage of lexical information. Second, I will describe a combined MEG-fMRI experiment aimed at localizing the effect of sentential context on lexical processing. As the results of this experiment were inconclusive, I then report a meta-analysis of previous imaging experiments that localize contextual effects. The results of this meta-analysis support the claim that context does impact activity in lexical storage areas, and therefore that predictive pre-activation of lexical representations is likely to play a role in normal sentence processing.

3.2 Neural basis of stored lexical-semantic representations

Semantic processing of linguistic material in the normal case (words organized into sentences in speech or print) must minimally involve activation and selection of candidate lexical representations and integration of the semantics of the selected representation with the context constructed on the basis of the previous words. Studies using fMRI, intracranial recordings, and neuropsychological phenomena have implicated three main regions as being involved in these computations: left inferior frontal cortex, left anterior temporal cortex, and left posterior temporal cortex.

Of these, the best candidate for the storage and access of amodal lexical representations is the region encompassing the left mid-posterior middle temporal gyrus (MTG) and parts of the neighboring superior temporal sulcus (STS) and inferior temporal cortex (IT) (Damasio, 1991; Indefrey & Levelt, 2004; Hickok & Poeppel, 2004; Hickok & Poeppel, 2007; Martin, 2007). The left STS and MTG show fMRI repetition priming (reduced BOLD activity on the second presentation) for auditorily presented words, but not pseudowords (Gagnepain et al., 2008); this priming is not observed when the task is non-lexical (phonological or first-letter/last-letter alphabetization; Gold et al., 2005). fMRI studies employing semantic tasks such as those requiring semantic categorization of words or judgments on their semantic properties consistently show activity in this region relative to other kinds of tasks (Price et al., 1994; Pugh et al., 1996; Cappa et al., 1998; Gitelman et al., 2005; Wagner et al., 2000), and studies using distorted speech stimuli find that activity in MTG/IT increases as a function of intelligibility (Davis & Johnsrude, 2003; Giraud et al., 2004). Increased activity in MTG is also observed when the number of words processed per trial is increased (Badre et al., 2005). Aphasia studies show that patients with lesions in posterior temporal areas have difficulty performing semantic tasks (Hart & Gordon, 1990; Kertesz, 1979). In particular, in a study of 64 patients, MTG was the only region in which lesions led to significantly lower performance on even the simplest sentences (Dronkers, Wilkins, Van Valin, Redfern, & Jaeger, 2004). With respect to language production, a meta-analysis of 82 imaging studies found that the left MTG was the only area reliably activated for tasks that required lexical selection (Indefrey et al., 2004).

These studies suggest that storage of lexical-semantic information is specific to the middle part of posterior temporal cortex—MTG and possibly STS and the inferior aspect of IT. More ventral parts of IT are associated with the representation of non-linguistic visual object features (Reddy & Kanwisher, 2006). The posterior superior temporal gyrus (STG) has sometimes been associated with semantic processing, but most existing evidence suggests that its role is limited to early (auditory) stages of the sound-to-meaning transformation (Binder et al., 2000) consistent with early models such as

Wernicke’s (Wernicke, 1874). Imaging studies typically show posterior STG activation for speech and other spectrotemporally complex stimuli regardless of semantic content (Wise et al., 2001). Although several fMRI studies show intelligibility effects in posterior STG as well as in MTG/IT (Davis & Johnsrude, 2003; Zekveld et al., 2006), it has been suggested that these effects may be driven by top-down processing (Scott, 2005; Davis & Johnsrude, 2007).

Work from various domains thus converges on the idea that left posterior middle temporal cortex subserves long-term storage and access of information associated with lexical representations. However, which aspects of lexical representation are stored here remains an open question. One view is that this region does not store semantic information per se, but rather stores the lexical representations that interface with a semantic network distributed across brain regions (Hickok & Poeppel, 2007), and that activation in this region in response to tasks involving pictures involves implicit lexical access (Vandenberghe et al., 1996). However, the results reviewed are also consistent with a view in which MTG stores some part of the semantic or conceptual features associated with lexical representations. Here I will remain agnostic on this question. For me, the important point is that this area is involved in storage of some kind of lexical information; therefore if the context affects this area’s response to a new input, it can be considered evidence that context influences the state of long-term memory representations rather than the processes involved in integrating the input into the larger, derived representations under construction.

Left inferior frontal cortex, left anterior temporal cortex, and recently left inferior parietal cortex have also been implicated in processing words in context, and have been variously attributed to mechanisms of lexical selection, lexical inhibition, semantic retrieval, thematic interpretation, syntactic composition, and semantic combination. I will review evidence on these regions in further detail in Chapter 4. For our purposes in this chapter, the important point is that these areas have mainly been argued to be involved in higher-level processing. Therefore, if contextual effects on lexical processing localize to one of these regions, it would lend support to a view according to which these effects reflect some aspect of integrating the input with the prior context, whether it is the process of selecting the representation to be entered into the derived structure, the specific computations required to integrate the current input, or an assessment of the well-formedness of the derived representation after it is constructed.

Finally, one caveat to this approach is that, for localization of contextual effects to a particular level of representation to constitute strong evidence for top-down mechanisms, it must be the case that the input could only be predicted with reference to a higher level of representation. For example, if it were the case that the sentence context contained individual words that were highly associated with the target word, reduction of activity at the lexical level could be instantiated by automatic spreading activation between lexical representations. The fact that the N400 context effect is observed when the broader sentence or discourse predicts the target even in the absence of lexical associations is what allows us to argue that if the effect localizes to a lexical storage area, it reflects a top-down mechanism.

In the experiment described in the next section we used a combined MEG and fMRI design in an attempt to determine whether the N400 effect of sentence context is truly top-down in nature by finding the neural source of the effect. This experiment represents joint work with Henny Yeung, Ryu Hashimoto, Allen Braun, and Colin Phillips.

3.3 Experiment 2: fMRI and MEG of sentence context effects

As I will detail in the next section, a number of studies have attempted to determine the neural source of the N400 effect of sentence context through intracranial recordings, ERP, MEG, and fMRI. Taken together, the results of these studies are inconclusive: the MEG studies most often localize the effect to mid-posterior temporal cortex, the fMRI studies to inferior frontal cortex, and the intracranial recordings to anterior medial temporal cortex.

These inconsistencies may be due to specific weaknesses of each method. Intracranial recordings on pre-surgical epileptic patients can only be recorded from regions that are clinically relevant, and therefore do not provide a complete picture of potential sources; MEG localization depends on the particular source localization algorithm chosen, and can be subject to significant error, especially if multiple sources show effects in the same time window; and fMRI has good precision in localization but cannot fix effects precisely in time. However, these inconsistencies may also be due to differences in materials, participants, or presentation parameters between studies. Differences in materials are a particular cause for concern, as many of the MEG and fMRI studies explicitly manipulated context-target congruity but may have varied in the strength of the lexical prediction engendered by the context.

In order to determine whether differences in materials, participants, or presentation parameters could be responsible for the differences in localization across techniques, we conducted a sentence context manipulation with the same materials, participants, and presentation parameters across both MEG and fMRI in two separate sessions. The timing information provided by the MEG data allowed us to determine when contextual effects occurred, and in particular allowed us to ensure that the materials we used in the fMRI session did indeed result in a N400 effect of sentence context.

Participants

Participants were 10 right-handed native speakers of English (6 male) with no known neurological disorders or abnormalities, each of whom participated in separate fMRI and MEG sessions. The sessions were separated by at least one week, and order of session was counterbalanced across participants.

Materials

Targets consisted of 13-word sentences containing a main clause with an object relative clause, as in (8). Three versions of each item were created: a control version, a semantic anomaly version, and a syntactic anomaly version. The syntactic anomaly condition was mainly included to address a different set of experimental questions, so I will not discuss the results for this condition in the thesis.

(8) Irene repaired the locker that the older schoolgirl…
   a. Control: …cruelly **<u>dented</u>** with a hammer.
   b. Semantic: …cruelly **<u>taunted</u>** with a hammer.
   c. Syntactic: …cruelly **<u>dent</u>** with a hammer.

Every target sentence followed exactly the same pattern as seen in (8): Name-verb-determiner-noun-‘that’-determiner-adjective-noun-adverb-verb-preposition-determiner-noun. The three sentences in a given set were identical except for the second, relative clause verb. In the control condition, the verb was semantically congruent with the context and correctly inflected for third-person past tense in agreement with the subject of the relative clause. In the syntactic anomaly condition, the verb was semantically congruent with the context, but it was missing the grammatically required inflection. In the semantic anomaly condition, the verb was correctly inflected for past tense, but it was semantically incongruent with the context. Specifically, the relative clause verb in this condition did not felicitously take the head of the relative clause as an object (e.g., in the example here, *taunt the locker*). Care was taken to select verbs that were felicitous but that were not strongly associated with the subject or object, based on experimenter intuition. Critical verbs were matched for frequency and length, and the same verbs were used in the correct and anomaly conditions. Different subsets of items were used in the fMRI and MEG sessions.

Note that, in contrast to the experiment presented in Chapter 2, the critical word in this experiment was always embedded within the sentence; that is, it was not sentence-final. Previous authors have argued that effects observed on sentence-final words conflate the response to the word with end-of-sentence wrap-up effects (e.g. Osterhout, 1997). By ensuring that the sentence continues several words after the target, we minimize the chance that the MEG response to the target will reflect sentence wrap-up effects, although it remains the case that the less temporally precise fMRI response is likely to reflect differential activity throughout the sentence.

In both sessions, a number of fillers equal to the number of experimental sentences were also included. These included both grammatical and syntactically/semantically anomalous fillers, in which the anomaly occurred in different positions of the sentence from the experimental targets so as to mitigate strategic processing to some extent. The number of anomalous sentences in the fillers was selected so as to make the experiment-wide ratio of good to anomalous sentences 1:1.

Procedure

Stimulus presentation parameters were matched across the MEG and fMRI recordings. Sentences were presented with central RSVP (rapid serial visual presentation), with an SOA (stimulus-onset asynchrony) of 500 ms (300 on, 200 off); thus the total sentence presentation time was 6.5 s. At the end of each sentence, an ACCEPT/REJECT? screen appeared for 2 s during which time the button-press response was collected. A 750 ms fixation screen followed all trials; thus the time allotted to each trial and its subsequent fixation was 9.25 s. Participants' task in both sessions was to judge each sentence as "acceptable" or not, where acceptability was explained to be dependent on both semantic and syntactic well-formedness. Participants were instructed to respond only when the sentence was completed and the ACCEPT – REJECT screen appeared. A brief practice session (~15 sentences) preceded the task. Responses were collected on a non-magnetic button box held in the right hand.

The fMRI session consisted of six 9-minute runs of 56 trials each. The 56 trials were composed of 24 target sentences (8 from each of the three experimental conditions), 24 filler sentences, and 8 null fixation trials to serve as the baseline and to improve the power of the deconvolution, for a total of 48 targets per condition across the 6 runs. In null trials, a fixation cross appeared onscreen for 6.5 s, followed by the response period; participants were instructed to choose the 'REJECT' button for these trials. Including set-up time and the anatomical and clinical scans that followed the experiment, fMRI sessions lasted ~2-2.5 hours.

The MEG session consisted of 18 blocks of 22 sentences each (11 target sentences and 11 filler sentences), for a total of 66 targets per condition across the session. Participants were allowed to pace themselves through the breaks. Including set-up time, MEG sessions lasted ~2.5 hours.

Recordings

MEG recordings were conducted using a 160-channel axial gradiometer whole-head system (Kanazawa Institute of Technology, Kanazawa, Japan). Subjects lay supine in a dimly lit magnetically shielded room (Yokogawa Industries, Tokyo, Japan) and were screened for MEG artifacts due to dental work or metal implants. A localizer scan was performed in order to verify the presence of identifiable MEG responses to 1 kHz pure tones (M100) and determine adequate head positioning inside the machine. The MEG signal was sampled at 500 Hz, and data were acquired continuously with an online bandpass filter of 1-200 Hz and a notch filter of 60 Hz.

fMRI data was recorded using a 3T General Electric system. Twenty-four axial slices (5 mm thickness, 1 mm inter-slice distance, FOV 19.2 cm, data matrix of 64 x 64 voxels) were acquired every 2 s using a BOLD sensitive gradient EPI sequence (TR = 2 s, TE = 30 ms, flip angle = 90 degrees). For each participant, a high-resolution 3D structural scan was acquired consisting of an MPRAGE sequence (124 sagittal slices, 1.2mm thickness, TR 7500 ms).

Analysis

An automatic noise reduction procedure was applied to the raw MEG data using three orthogonal magnetometers as reference with a time-shift PCA filter (de Cheveigné

& Simon, 2007). Epochs of 1100 ms (including a 100 ms pre-stimulus baseline) were collected at the critical verb region in each comparison. Epochs with ocular and motion artifacts exceeding 2 pT in amplitude were identified by visual inspection and removed before averaging. For the analyses presented below, data were averaged for each condition in each participant, and baseline corrected using the 100 ms pre-stimulus interval. For the figures, a low-pass first-order half-power cutoff elliptical filter was used for smoothing (default ripple in passband of 3 db; default stopband attenuation of 50 db). All statistics presented were computed prior to offline filtering.

The small sample size of this study precluded us from selecting channels of interest by creating a statistically thresholded topographical map as we did for the experiment presented in Chapter 2. Therefore, we began by examining the RMS of activity averaged across all sensors in each hemisphere for the control and semantic anomaly conditions to obtain a coarse estimate of the time course of the MEG response to syntactic and semantic anomalies without building in assumptions about localization. From the topography of the response during the N400 time-window we determined that examining the average activity for each quadrant, as we did in Chapter 2, would provide a reasonable measure of differential activity.

In order to assess the approximate source of the differential activity in the N400 time-window, we fit equivalent current dipoles (ECDs) for the six participants who showed a clear dipole pattern in the left hemisphere in this time-window. Based on visual inspection of the scalp distribution in the left hemisphere for each participant at the largest peak in the N400 time-window for the critical semantic anomaly condition, we chose 5 channels from the sink and 5 from the source that best represented the dipolar distribution. We then used these channels to fit an ECD for each participant using MEG160 software. The shape of the conducting volume was modeled as a sphere defined on the basis of each participant's head shape data. Difficulties in image coregistration have so far precluded us from using the structural MR image to model the shape of the conducting volume.

Data analysis was performed on fMRI data with the AFNI software package (Cox, 1996). Functional images were corrected for temporal and spatial drift, a spatial filter of 8 mm was applied, and the resulting data were normalized. Individual coefficient matrices were derived through a fixed-HRF deconvolution, with the null fixation trials serving as the baseline. We followed Kuperberg and colleagues (2003) in splitting the duration of each target sentence into two parts (the first six words from the last seven words), and contrasting the response across conditions to the second part of the sentence where the critical word that varied across conditions (the 10th word) was presented. The filler items were also included as part of the model, modeled separately as normal, semantically anomalous, and syntactically anomalous fillers. Because the anomaly occurred in variable position in the filler items, the duration of the entire filler sentence was modeled together. The coefficient matrices were entered into a within-subjects ANOVA in which main effects of condition and contrasts were computed. All fMRI images are presented in neurological convention (L=R).

MEG Results

Figure 12 shows the RMS of activity averaged across all sensors in each hemisphere for the control and semantic anomaly conditions. Visual inspection indicates a robust difference in amplitude between the two conditions in the N400 time-window

(300-500 ms) in the left hemisphere only, followed by a bilateral difference in the time-window associated with the late positivity or P600 (600-900 ms). Paired sample t-tests indicated that these RMS differences were significant in the N400 window in the left hemisphere only (LH - $p < .05$; RH - $p > .1$) as well as in the P600 window bilaterally (LH - $p < .01$; RH - $p < .05$).

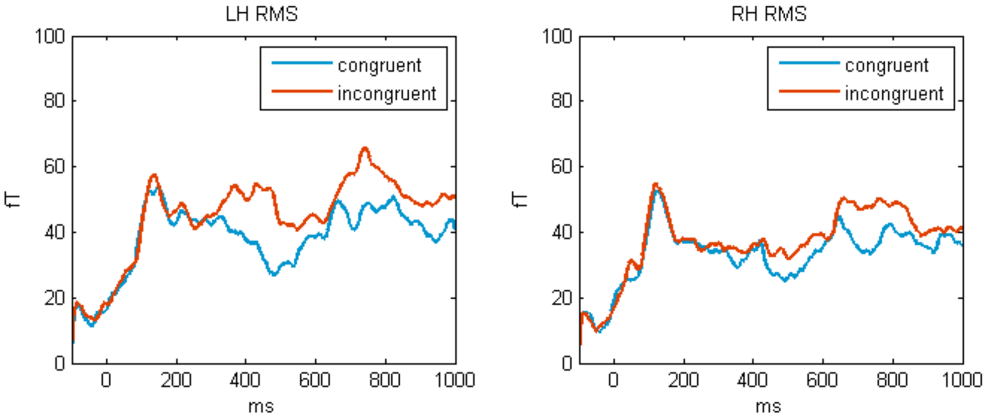
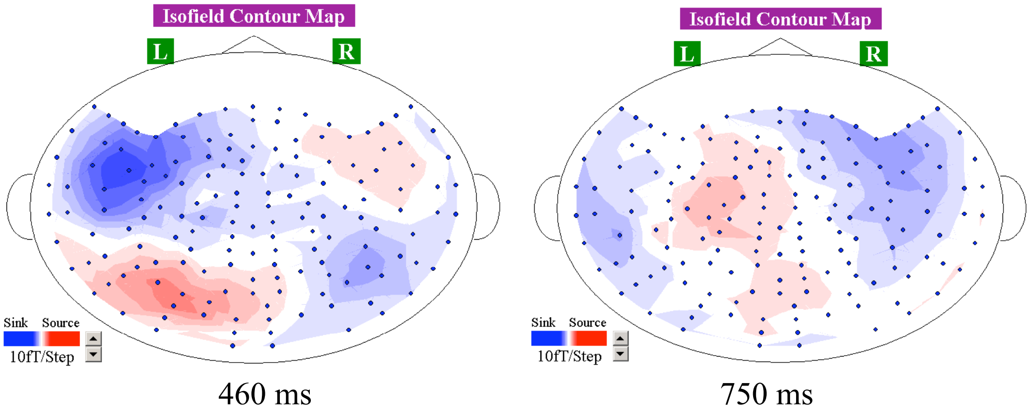


Figure 12. RMS of activity averaged across all sensors in each hemisphere for the control and semantic anomaly conditions.

Figure 13 shows the grand-average topography of the difference (incongruent-congruent) at each of these time-windows. During the N400 time-window, we observe a difference map similar to that observed for the same comparison in Chapter 2 in the left hemisphere: a left anterior cluster of sensors with a negative sign, and a left posterior cluster of sensors with a positive sign. In contrast to Experiment 1, however, the right anterior cluster of sensors showing a positive sign shows only a weak difference. During the P600 time-window, the difference map also shows a weaker effect, which does not resemble the difference map for the late positivity observed in Chapter 2, nor any other MEG responses that I am aware of.



460 ms

750 ms

Figure 13. Grand-average difference maps for incongruent-congruent comparison at target word. Activity is shown at the approximate peak of the difference wave in each window: 460 ms for the N400 time-window and 750 ms for the late positivity time-window.

The pattern observed in the grand-average difference map during the N400 time-window suggests that splitting the sensor space into quadrants should approximately separate sensors that showed effects of different polarity in the same hemisphere. This allows us to plot the average over sensors rather than the RMS, which avoids potential distortion introduced by RMS and which allows us to make a closer comparison to the results of Experiment 1.

Figure 14 presents the raw averages and the difference waves across the sensors divided into quadrants. These figures confirm that, across quadrants, a robust effect is observed between 300-500 ms in left posterior and left anterior sensors, but not in right hemisphere sensors. Although the main goal of this experiment was to compare the putative source of the N400 effect across techniques, in the Discussion I examine the question of why the MEG results were slightly different from those in Chapter 2.

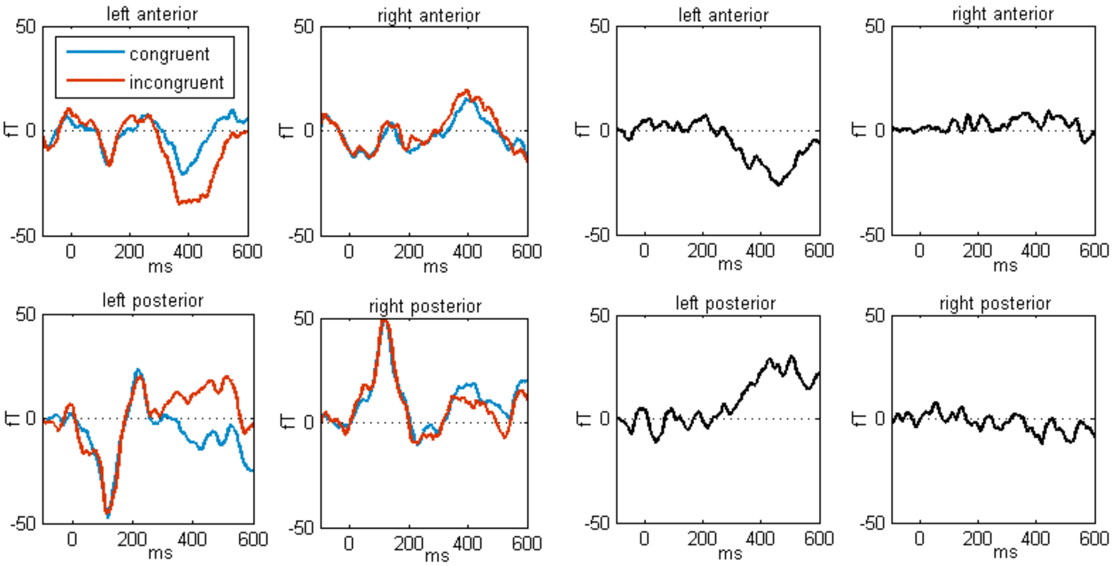


Figure 14. On the left, average activity across MEG sensors divided into four quadrants for the control and the semantic anomaly conditions; on the right, the difference waves resulting from subtracting the incongruent-congruent waves in each quadrant.

Figure 15 presents the equivalent current dipole fits for the semantic anomaly condition for each of the six participants who showed a left hemisphere dipole pattern. The best-fitting dipole for each of the six participants localized to a position consistent with left mid-posterior temporal cortex. Despite the anatomical imprecision of source localization using a head shape model, the results clearly implicate a somewhat posterior source and not a frontal or anterior temporal source, and are consistent with other MEG studies that used structural MRIs for ECD source modeling (e.g. Halgren et al., 2002).

fMRI Results

The fMRI results presented in Figure 16 and

show the areas that were active for all sentences relative to fixation. This contrast showed activity across all the classical language areas: left inferior frontal cortex, left temporal cortex from the anterior to the posterior extent, and parts of left inferior parietal cortex including angular gyrus.

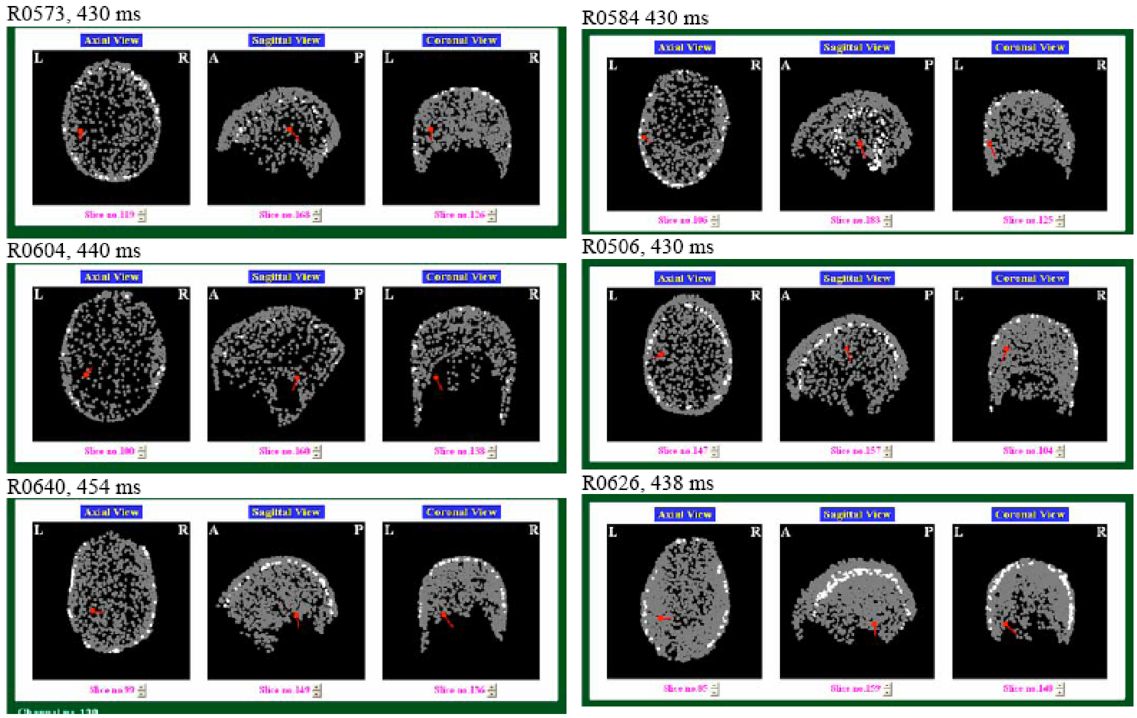
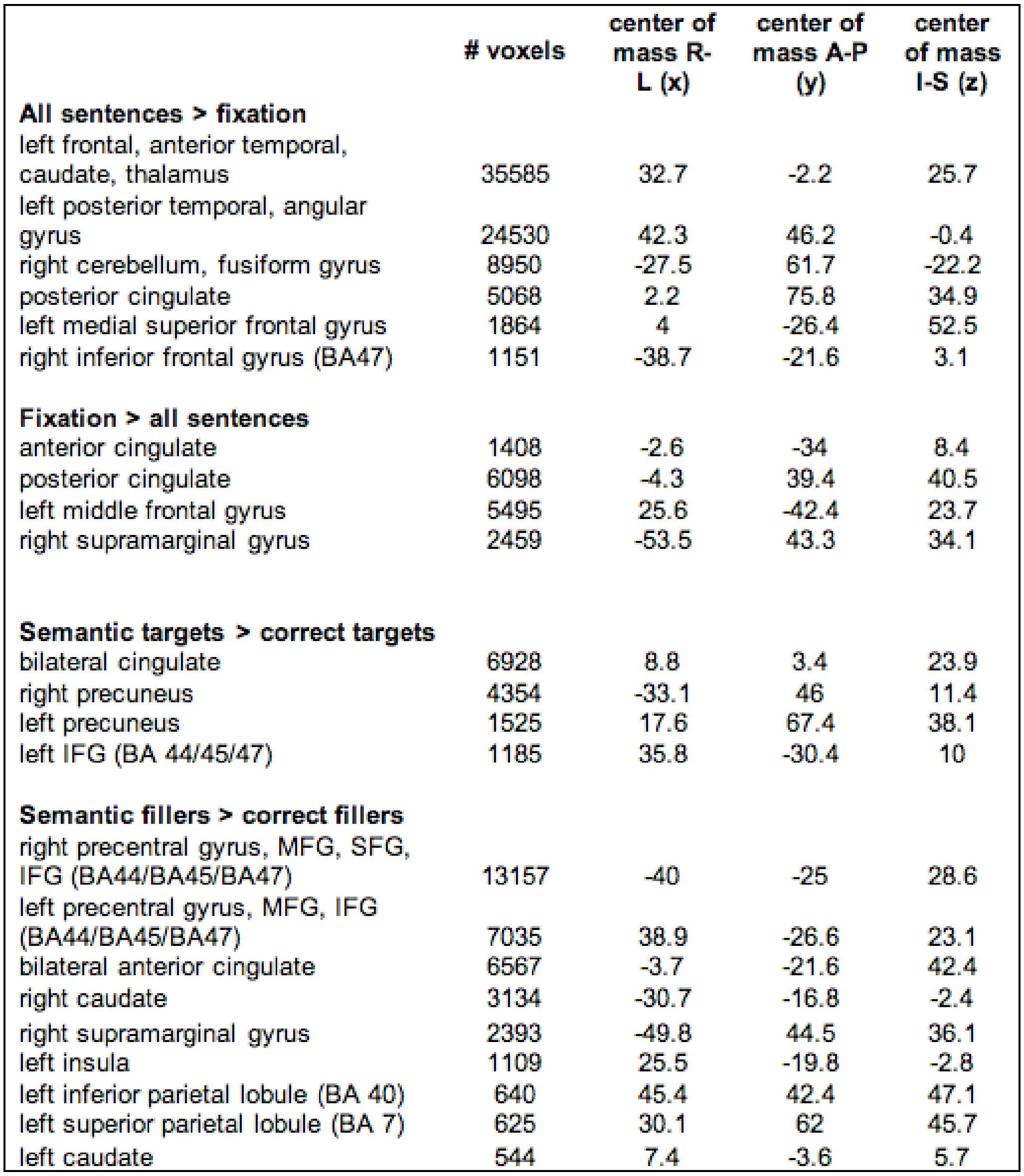


Figure 15. Digitized head shape dipole fits for the semantic anomaly condition for six participants. The dipole was modeled at the peak of the N400 component for each participant; the latency of this peak is indicated in each case.

Figure 16 and also show the areas that were active for the second half of the semantic anomaly targets (the part of the sentence including the critical word 10) relative to the second half of the correct targets. These areas were an area of left precentral gyrus that included part of posterior inferior frontal cortex (BA 44), left caudate, and a mid-posterior region of left inferior frontal cortex (BA 44/45).

Because our fMRI model included a separate factor for the semantic anomaly filler items (fillers that included semantic anomalies but were not controlled for sentence structure or type or position of anomaly), we also examined the contrast between semantic anomaly fillers and correct fillers. This contrast showed significant differences in a number of areas, including a larger area of left inferior frontal cortex (BA 44/45/47) and a corresponding area of right inferior frontal cortex.



	# voxels	center of mass R-L (x)	center of mass A-P (y)	center of mass I-S (z)
All sentences > fixation				
left frontal, anterior temporal, caudate, thalamus	35585	32.7	-2.2	25.7
left posterior temporal, angular gyrus	24530	42.3	46.2	-0.4
right cerebellum, fusiform gyrus	8950	-27.5	61.7	-22.2
posterior cingulate	5068	2.2	75.8	34.9
left medial superior frontal gyrus	1864	4	-26.4	52.5
right inferior frontal gyrus (BA47)	1151	-38.7	-21.6	3.1
Fixation > all sentences				
anterior cingulate	1408	-2.6	-34	8.4
posterior cingulate	6098	-4.3	39.4	40.5
left middle frontal gyrus	5495	25.6	-42.4	23.7
right supramarginal gyrus	2459	-53.5	43.3	34.1
Semantic targets > correct targets				
bilateral cingulate	6928	8.8	3.4	23.9
right precuneus	4354	-33.1	46	11.4
left precuneus	1525	17.6	67.4	38.1
left IFG (BA 44/45/47)	1185	35.8	-30.4	10
Semantic fillers > correct fillers				
right precentral gyrus, MFG, SFG, IFG (BA44/BA45/BA47)	13157	-40	-25	28.6
left precentral gyrus, MFG, IFG (BA44/BA45/BA47)	7035	38.9	-26.6	23.1
bilateral anterior cingulate	6567	-3.7	-21.6	42.4
right caudate	3134	-30.7	-16.8	-2.4
right supramarginal gyrus	2393	-49.8	44.5	36.1
left insula	1109	25.5	-19.8	-2.8
left inferior parietal lobule (BA 40)	640	45.4	42.4	47.1
left superior parietal lobule (BA 7)	625	30.1	62	45.7
left caudate	544	7.4	-3.6	5.7

Table 2. Brain regions that reached a significance level of $p < .01$, uncorrected for the contrasts indicated. Only clusters in which greater than 500 adjacent voxels showed a significant contrast are included. No regions reached this criterion for the contrasts *correct targets > semantically anomalous targets* or *correct fillers > semantically anomalous fillers.* IFG = inferior frontal gyrus, SFG = superior frontal gyrus, MFG = middle frontal gyrus.

No regions showed significantly more activity in non-anomalous target or fillers relative to semantically anomalous targets and fillers. Other regions were also associated with greater activity in the semantically anomalous targets and fillers, including the anterior and posterior cingulate, superior and middle frontal gyri, and precentral gyrus bilaterally. Because the previous literature has not proposed functions for these areas specific to language processing, I will not speculate on their role here, but they may reflect differences in attention or error-detection mechanisms between the anomalous and control materials.

Figure 16. fMRI contrasts of interest from Experiment 2 in the left hemisphere. All images display voxels that demonstrated a significant difference between conditions across 10 participants, $p < .01$ uncorrected. Orange indicates that the sign of the contrast was positive and blue indicates negative. The left image illustrates the contrast of language comprehension - visual fixation; the central image illustrates the contrast of second-half semantically anomalous experimental targets – second-half control experimental targets; the right image illustrates the contrast of semantically anomalous fillers – correct fillers.

Discussion

As in the previous literature, in this experiment MEG and fMRI seemed to implicate different regions of cortex in responding differentially to congruent, predictable endings and incongruent, unpredictable endings. The MEG field pattern observed for the difference between incongruent and congruent sentence endings was similar to the field pattern observed for other MEG components thought to be generated in temporal cortex (e.g., the M100 auditory response), and equivalent current dipole localizations during the N400 time-window for six participants localized dipoles to mid-posterior temporal areas, replicating previous studies (e.g. Helenius et al., 1998; Halgren et al., 2002; Pylkkänen &

McElree, 2007). On the other hand, the fMRI results indicated significant differences in the response to the same incongruent and congruent targets in inferior frontal cortex, and did not reveal significant differences in any temporal areas for this contrast, also replicating previous studies (e.g. Friederici et al., 2003; Hagoort et al., 2004). The MEG results, in localizing the effect to posterior temporal cortex, would support a top-down mechanism for sentence context effects, while the fMRI results would not. This contrast was observed even though the participants, procedure, and items that were used across the two measures were identical.

Why did MEG demonstrate a posterior temporal effect while fMRI did not? One possibility is that invalid assumptions of the equivalent current dipole model used here—e.g. that there is just a single dipole in the left hemisphere at the critical time-point and that this dipole is represented as a single point—resulted in an inferior frontal source being mistakenly displaced to posterior temporal cortex. Although this is a possibility, I am not aware of any known examples of this kind of displacement happening with ECD models of MEG data.

Another possibility is that the fMRI failed to find a real effect in posterior temporal cortex. One piece of supporting evidence for this possibility is that the contrast between semantically anomalous fillers and correct fillers found a small but significant effect in left posterior inferior temporal cortex (we could not examine the corresponding MEG effect because the point at which the sentences became anomalous was not time-locked in the fillers). If we assume that the processes engaged by the fillers were similar to those engaged by the more tightly controlled targets, this result could be considered evidence for a posterior temporal locus for the effect of contextual support. It is still not clear, however, why the same contrast in the targets would fail to show the same effect. One account that I discuss in more detail in the meta-analysis in the next section is that the longer time-scale of the fMRI measurement makes it more difficult to pull out differences in lexical storage areas in sentence experiments. The idea is that because processing each word of the sentence will involve activation of these areas, and because other parameters like lexical frequency will modulate the activation in these areas on each word, there is a background of continuous variation in activity in this area across the sentence. Against this background of variation it might be harder for activity resulting from the critical word in this area to show significant differences.

One difference between the semantically anomalous fillers and the targets was that in the fillers the anomaly could come much earlier in the sentence. On a prediction account, this could result in not only the anomalous word being unpredictable, but all the subsequent words in the sentence being less predictable. This in turn would result in a longer-lasting difference in posterior temporal areas that would have a better chance of surviving the surrounding variation, and could therefore explain why the fillers demonstrated a significant posterior inferior temporal effect and the targets did not.

Even if we assume this account to be correct, it remains to be explained why the robust inferior frontal effect in both targets and fillers observed in fMRI was not observed in MEG. One possible explanation is that the frontal effect did not generate the MEG effect observed in the N400 time-window, but rather generated the significant difference observed in the RMS across sensors in both hemispheres in the later, 600-900 ms time-window associated with late positivities in ERP. However, because the field pattern for the contrast was weak in this time-window and did not resemble a typical dipole pattern, the MEG data does not provide strong evidence for this conclusion.

Other accounts of the MEG and fMRI discrepancy we observe here are also possible. For example, the MEG localizations to temporal cortex may be slightly more anterior than they appear. If the effect of context is due to differential activity in anterior temporal cortex, the lack of an effect in fMRI may be due to known susceptibility to artifact in this region, discussed in more detail in the next section.

In sum, the MEG data described here, in demonstrating a posterior temporal source for contextual effects on lexical processing, support a top-down mechanism for these effects, but the fMRI data, which have more spatial precision than MEG data, do not provide clear evidence for a posterior temporal source. This discrepancy replicates the same discrepancy seen between most previous MEG and fMRI studies of sentence context effects (Van Petten & Luka, 2006), and makes it difficult to use localization as an argument for the top-down nature of contextual effects. In the next section, I examine in more detail previous efforts to localize contextual effects. I will conclude that, taken as a whole, this literature supports a posterior temporal source for contextual effects, and thus supports the claim that a top-down mechanism is responsible for such effects.

Finally, I note as an aside that we observed several interesting differences in the MEG response to the contextual manipulation in Experiment 2 as compared to Experiment 1. When we applied the same quadrant analysis to the data from Experiment 1, we found that, in addition to the two left hemisphere quadrants, the right anterior quadrant showed a significant difference in the 300-500 ms time-window between contextually supported and unsupported conditions ($t(17) = 3.05$, $p < .01$) in contrast to

Experiment 2 in which only the two left hemisphere quadrants showed a significant effect. This right hemisphere difference appeared to be driven by a reduction in activity for the supported condition in Experiment 1 that was not observed in Experiment 2. We also observed that the peak latency for the difference wave appeared to be earlier in Experiment 1 (~350-400 ms) than in Experiment 2 (~450-500 ms). These differences could be due to a number of differences in materials and design: more participants and items were tested in Experiment 1, giving this design more power to detect differences; targets in Experiment 1 were sentence-final while in Experiment 2 they were sentence-internal; the task in Experiment 1 was a probe task unrelated to acceptability while the task in Experiment 2 was to judge acceptability; the anomalies in Experiment 2 required filling in an argument from earlier in the sentence when the verb was encountered; and the anomalies themselves tended to be more severe in Experiment 1, in some cases making any kind of semantic composition difficult (*Three people were killed in a major highway pumpkin*) while the anomalies in Experiment 2 tended to be composable violations of world knowledge (*the locker that the girl taunted*). Testing these differences more systematically in future experiments may provide a means for functionally dissociating the left-hemisphere effect observed in both experiments from the right-hemisphere effect observed only in Experiment 1.

3.4 Meta-analysis of previous results

I began this chapter with the goal of finding the cortical areas sensitive to the context in which a word is presented, in order to test a particular hypothesis of predictive effects in comprehension: that information flows directly from levels representing the broader context to the level at which lexical representations are stored and accessed. To test this hypothesis, we tested the same contextual manipulation in the same participants with two neurophysiological measures, fMRI and MEG. Unfortunately, we found that the two methods did not converge on the same areas: while MEG showed differential activity in the area associated with storage of lexical information, fMRI did not.

This non-convergence is representative of the pattern of results across a number of studies that have attempted to localize effects of context on lexical processing (Van Petten & Luka, 2006), which most frequently find effects in either inferior frontal or temporal areas. The results of Experiment 2 suggest that differences in participants, procedure, and materials cannot account for this variability. Therefore, I now return to make a more careful examination of the previous literature on contextual effects in comprehension to better understand this puzzle. In this chapter my main focus remains to determine whether or not the context reliably affects activity in the cortical area involved in storage of lexical information, posterior temporal cortex, because this will provide strong evidence for a particular kind of top-down/predictive mechanism in lexical processing. The role of inferior frontal cortex in language comprehension is less well understood and thus the effects found in this area do not so unambiguously constrain the functional interpretation of contextual effects. In Chapter 4, based on work in other domains, I will present a model in which anterior inferior frontal cortex mediates targeted retrieval of lexical/conceptual information based on context and in which mid-inferior frontal cortex mediates selection among multiple activated representations.

In this section I will review previous work that attempted to localize contextual effects on language comprehension in semantic priming and sentence context paradigms, across a number of techniques including ERP, fMRI, MEG, and intracranial recordings.

Because these paradigms elicit the N400 effect in ERP, the cortical areas reliably associated with contextual effects on lexical processing are likely to be those that are responsible for the observed differences in N400 amplitude.

ERP studies

The N400 context effect in ERP tends to have a centroparietal scalp distribution, with a small but consistent bias to the right side of the head when visual presentation is used (Kutas, Van Petten, & Besson, 1988). However, ERP studies with split-brain patients support a left-hemisphere generator for the N400 effect: only sentence completions presented to the 'linguistic' hemisphere elicited N400 effects (Kutas, Hillyard, & Gazzaniga, 1988). Given the strong left-lateralization for language observed elsewhere, the central-right scalp distribution of the N400 effect has been interpreted as "paradoxical lateralization" in which a left-hemisphere generator impacts right-hemisphere electrodes due to fissural morphology and conductance properties (Van Petten & Rheinfelder, 1995; Van Petten & Luka, 2006; Hagoort, 2008). Alternatively, the asymmetry could be due to an overlapping left-lateralized positivity for visual presentation, which would result in a smaller negativity in the waveform. Localization from ERP data has occasionally been attempted for the N400 effect (Curran, Tucker, Kutas, & Posner, 1993; Johnson & Hamm, 2000; Frishkoff et al., 2004), but results have been inconsistent.

Studies of patients with various forms of brain lesion can inform us about the generators of the N400 by showing that brain damage to particular regions alters the N400 effect. However, the existing data provide limited evidence. The effects of semantic congruity on the N400 are relatively preserved in patients with amnesia and

Alzheimer's disease (Iragui, Kutas, & Salmon, 1998; Olichney et al., 2002; Olichney et al., 2000), but this result is difficult to interpret since the full extent of areas affected by these disorders is unclear. Similarly, studies of patients with aphasia show that poor language comprehension is associated with a reduced N400 effect for priming and semantic anomaly, but the areas of damage in these patients are often unknown, and patients with Broca's aphasia and Wernicke's aphasia can show robust N400 effects (Hagoort, Brown, & Swaab, 1996; Swaab, Brown, & Hagoort, 1997; Kojima & Kaga, 2003). Of studies with smaller sample sizes, one showed N400 congruity effects in three patients with left frontal lesions (although these effects were attenuated; Friederici, von Cramon, & Kotz, 1999), while another found no N400 congruity effect in a patient with a left temporal lesion (Friederici, Hahne, & von Cramon, 1998). Some further evidence points to temporal lobe involvement, as patients with left temporal lobe epilepsy show no N400 congruity effect, in contrast to patients with right temporal lobe epilepsy (Olichney et al., 2002). In general, however, small sample sizes and heterogeneous etiologies make it difficult to associate damage to particular regions with presence or absence of N400 effects in existing patient studies.

fMRI studies

The use of fMRI results to provide evidence on the source of ERP effects faces several difficulties. The fMRI signal is much delayed relative to ERPs, does not reflect properties such as phase coherence that may contribute to ERP amplitude and may have lower signal-to-noise ratio in some cortical areas. These factors also constrain experimental designs. However, the spatial accuracy and wide availability of fMRI has led to a wealth of studies using the same contextual manipulations that are used to elicit the N400 effect.

Table 3 and Figure 17 summarize the results of 9 studies comparing fMRI signals in response to semantically related and unrelated word pairs – the semantic priming paradigm. The parameters used differ in modality, type of semantic relation, and crucially, stimulus-onset asynchrony (SOA)—the duration of the interval between the presentation of the prime and the presentation of the target.

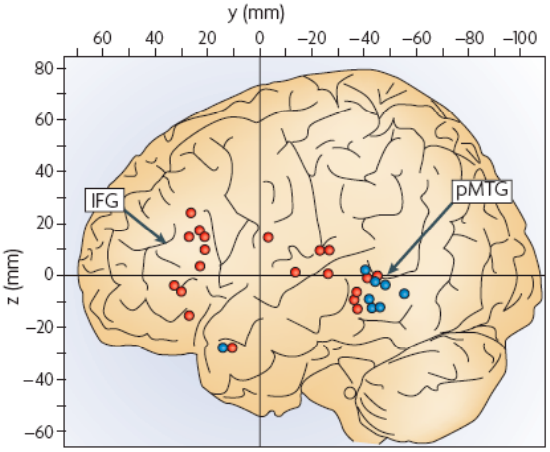


Figure 17. A visual summary of the results of semantic-priming manipulations in functional MRI. Approximate locations of centers of significant activation in the left inferior frontal and temporal cortices. These studies differed in stimulus-onset asynchrony (SOA); activation in short-SOA studies (< 250 ms) is shown in blue; activation in long-SOA studies (> 600 ms) is shown in red. Both short- and long-SOA studies found posterior middle temporal effects (mainly in MTG, with additional superior temporal effects in auditory studies), but only long-SOA studies showed inferior frontal effects.

Two patterns are evident across the priming studies in fMRI. First, the only area that, like N400 amplitude, consistently shows effects of priming across all different experimental parameters is left posterior MTG. Second, a subset of studies also showed left inferior frontal effects, but these effects were strongly dependent on SOA: only priming at long SOAs (> 600 ms) affected inferior frontal activity. These patterns were observed both across- and within-subjects (Gold et al., 2006).

Reference	Mod.	Task	SOA		IFG		IT	MTG	STG	TP	AG
				BA47	BA45	BA44	BA37	BA21	BA22	BA38	BA39
Short SOA studies											
Devlin et al., 2004	V	lexical decision	33					●			
Gold et al., 2006	V	lexical decision	150				●	●			
Rossell et al., 2003	V	lexical decision	200							●	
Gold et al., 2006	V	lexical decision	250				●	●			
Wheatley et al., 2005	V	naming target	250				●				
Long SOA studies											
Matsumoto et al., 2005	V	lexical decision	600	●	●						
Rissman et al., 2003	A	lexical decision	600					●	●		
Kotz et al., 2002	A	lexical decision	650		●	●			●		
Wible et al., 2006	A	lexical decision	750		●			●	●		
Giesbrecht et al., 2004	V	relatedness	800	●	●			●			
Gold et al., 2006	V	lexical decision	1000	●	●	●		●			
Gold et al., 2006	V	lexical decision	1000	●	●	●		●			
Rossell et al., 2003	V	lexical decision	1000							●	

Table 3. Significant effects for whole-head contrasts of primed and unprimed targets. Effects in right-hemisphere regions were relatively few and inconsistent across studies, and therefore only left-hemisphere effects are reported here. Effects in the temporal, the inferior frontal, and the inferior parietal cortices are reported; no other regions showed consistent effects across studies. Modality of presentation (auditory (A) or visual (V)), task, and stimulus-onset asynchrony (in milliseconds) are indicated for each contrast. AG, angular gyrus, BA, Brodmann's area; IFG, inferior frontal gyrus, IT, inferior temporal cortex; MTG, middle temporal gyrus; STG, superior temporal gyrus; TP, temporal pole.

At first glance, these findings appear to straightforwardly support the hypothesis that top-down/predictive mechanisms impact activity at lower levels of representation, as posterior MTG is thought to support lexical storage and access. However, one could argue that word pair semantic priming is due solely to direct connections between individual lexical representations, and thus that the reduction in activity observed in left posterior MTG in these studies does not reflect a top-down mechanism at all. Previous work on the interaction between SOA and priming provides some argument against this objection.

A great deal of behavioral research has supported a distinction between two kinds of priming (Posner & Snyder, 1975): automatic (in which activation automatically spreads through a network of representations that are semantically related to the prime) and strategic (in which the prime is used to generate expectancies for the target or to inform other heuristics for responding to the task quickly). At short SOAs, automatic spreading activation seems to dominate, presumably because there is not enough time for expectancies to be generated, and strategic processes are thought to play a role only at longer SOAs (Neely, 1977; Neely, Keefe, & Ross, 1989). However, N400 effects of similar size, timing, and topographical distribution are seen at both short and long SOAs in ERP studies (Anderson & Holcomb, 1995; Deacon et al., 1999; Hill et al., 2002; Franklin et al., 2007; Rossell, Price, & Nobre, 2003; Nakao & Miyatani, 2007), suggesting that the N400 effect reflects a process common to both (Figure 18).

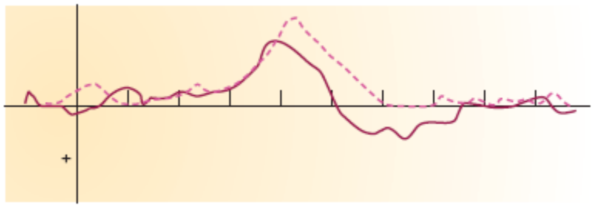


Figure 18. Event-related potential (ERP) difference waves (the waveforms that are obtained by subtracting the related from the unrelated condition) for semantic priming with 200 ms SOA (dotted line) and 800 ms SOA (solid line), figure modified from Anderson and Holcomb (1995). An N400 effect of similar amplitude is observed at both SOAs.

MTG was the only area that, like the N400 effect, demonstrated a reduction in activity across all SOAs, which strongly suggests that differential activity in MTG is the source of the differences observed in N400 amplitude. One possibility is that these differences reflect facilitation in MTG by both spreading activation (at short SOAs) and top-down expectancies (at long SOAs). However, another possibility is that these differences reflect spreading activation at both SOAs, and that top-down expectancies exert their effects in another way. The critical evidence against this second possibility comes from several ERP studies that pit spreading activation and top-down expectancies against each other using a classic paradigm from Neely (1977). In these studies, participants are told ahead of time that certain items from one semantic category (e.g., birds) are likely to be followed by items from a second, arbitrarily chosen semantic category (e.g. buildings). Then the response to an expected but unrelated target can be compared to the response to an unexpected related target. Behavioral studies observe only relatedness priming effects at short SOAs, and observe expectedness effects at long SOAs only. The same pattern is observed for the N400 effect in ERP: at short SOAs, an N400 amplitude is affected by relatedness but not expectancy, but at long SOAs, N400 amplitude is affected by both (Deacon et al., 1999; Nakao & Miyatani, 2007). Importantly, the N400 effect of expectancy is identical in timing and distribution to the N400 effect of relatedness. Priming effects due to expectancy in this paradigm are unlikely to be due to automatic spreading activation within the lexical level, because there are not long-term stored connections between the prime and the target in this condition; if there were, we would expect to see priming effects in the short SOA condition as well. Therefore, as the fMRI data suggests that the N400 priming effect is generated by MTG, as these ERP studies suggest that the N400 priming effect is associated with priming due to both spreading activation and top-down expectancies, and as behavioral studies suggest that top-down expectancies generally affect reaction times at long SOAs, there is a strong basis for assuming that the MTG effects observed in the long SOA priming studies were partially due to top-down mechanisms.

The SOA-dependence of frontal effects can be explained under a number of models that suggest inferior frontal cortex is involved in prediction, working memory, or strategic processes, if it is assumed that the differential frontal activity is somehow related to the generation of expectancies, prime-target matching, or other mechanisms available when there is more time between the prime and target. According to the model I present in Chapter 4, the long SOA allows the system to generate predictions that may either facilitate the controlled retrieval process initiated by anterior LIFG in related pairs or increase the burden on the selection process initiated by posterior LIFG in unrelated pairs. On the other hand, the absence of frontal effects in short SOA studies reflects insufficient time for the prime to be processed and entered in the current context, so that the prime cannot influence frontal processes of selection and controlled retrieval based on contextual information.

Some past studies have implicated a larger part of the posterior temporal cortex (including the STS and superior IT) in the storage of lexical information. However, these areas did not show up consistently in the fMRI semantic priming paradigms shown in Table 3. These data show that activation in IT and STG is sometimes observed, but is dependent on specific design parameters of the priming studies (short SOA for IT and auditory presentation for STG), suggesting that their absence in other conditions is not simply a result of spatial smoothing in the fMRI analysis.

Ref	Mod.	Task	IFG			FFG	IT	MTG	STG	TP	AG
			BA47	BA45	BA44	BA36	BA37	BA21	BA22	BA38	BA39
Ni et al., 2000	A	animacy			●				●		
Friederici et al., 2003	A	accuracy judgment							●		
Kuperberg et al., 2003	V	accuracy judgment	●	●	●			●	●		
Kiehl et al., 2002	V	accuracy judgment	●	●		●	●			●	
Kuperberg et al., 2008	V	accuracy judgment	●			●					
Ferstl et al., 2005	A	accuracy judgment	●								
Rüschemeyer et al., 2005	A	accuracy judgment	●								
Rüschemeyer et al., 2005	A	accuracy judgment		●							
Rüschemeyer et al., 2005	A	accuracy judgment	●	●				●			
Rüschemeyer et al., 2006	V	accuracy judgment			●						
Stringaris et al., 2007	V	accuracy judgment	●				●				●
Newman et al., 2001	V	accuracy judgment									●
Cardillo et al., 2004	A	lexical decision		●	●						
Baumgaertner et al., 2002	V	lexical decision	●								
Hagoort et al., 2004	V	none	●	●							
Kang et al., 1999*	V	none		●	●						
Dien et al., 2008	V	none		●							

Table 4. Significant effects for whole-head contrasts of anomalous and congruous sentences in left-hemisphere language areas. Effects in right-hemisphere regions were relatively few and inconsistent across studies, and therefore only left hemisphere effects are reported here. Effects in the temporal, the inferior frontal and the inferior parietal cortices are reported; no other regions showed consistent effects across studies. The modality of presentation (auditory (A) or visual (V)) and task are indicated for each contrast. AG, angular gyrus; BA, Brodmann's area; FFG, fusiform gyrus; IFG, inferior frontal gyrus; IT, inferior temporal cortex; MTG, middle temporal gyrus; STG, superior temporal gyrus; TP, temporal pole. *This study used verb phrases (for example, 'broke rules' and 'ate suitcases') rather than full sentences.

fMRI results for sentence-level studies are less consistent. Table 4 summarizes the results of 15 studies that contrasted activity in response to sentences with normal and semantically implausible/unexpected endings. The only region in which activity was consistently affected by contextual semantic fit was the left inferior frontal cortex. These effects were seen in both anterior and posterior inferior frontal areas, although this differed across studies. Only 6 cases showed a significant effect in left temporal regions, and the location of these effects was highly variable. Note that in sentential contexts SOA is not a relevant factor, because predictive information accumulates gradually rather than being made available at a single point; therefore the duration of the interval between the pre-critical word and the critical word should have much less of an impact on the availability of the prediction.

Previous authors have attributed the effects of sentential context on LIFG activity to strategic, task-related processes rather than mechanisms normally engaged in language processing (Van Petten & Luka, 2006). However, LIFG effects were robust across diverse tasks, and were found even in passive reading paradigms. In the next chapter, I will review previous work that suggests that LIFG underlies normal processes of targeted retrieval and selection of lexical representations in anterior and middle IFG, respectively. This hypothesis can account for the LIFG effects observed in sentence context manipulations and in long SOA priming experiments. The electrophysiological correlates of the LIFG effects are less clear. In the next chapter I discuss several possibilities.

More surprising is the inconsistency of the effects of sentential context in temporal cortex. This is problematic, given that MTG is the only region that matches the N400 pattern in showing semantic priming effects across SOAs. I suggest that MTG effects have failed to show up in existing fMRI results for several reasons. First, the temporal insensitivity of fMRI means that it is most effective at identifying distinctive activation within an interval of a few seconds. Modulation of lexical access processes in priming contexts is fairly distinct, as the prior context consists only of a single word; however, modulation of lexical access due to a predictive sentence context may not stand out against the many and varied lexical access events that occur over the course of any sentence. Distinguishing lexical access effects from this background may require more targeted fMRI designs. Second, the manipulations of sentential context used in most fMRI studies differed from those used in most ERP N400 studies, in which the sentence contexts were designed to strongly predict a particular ending such that the congruent conditions are also predictable conditions. Conditions that differ only in semantic congruity show smaller differences in N400 amplitude (e.g., Kutas & Hillyard, 1984; Connolly & Phillips, 1994), consistent with the view that the 'semantic anomaly' N400 effect is driven not by anomaly primarily but by predictive pre-activation in the congruent condition. In many of the fMRI studies, materials were designed to be anomalous but not necessarily predictive, and thus may have elicited weaker effects.

According to the view that the N400 context effect is due to the facilitative effects of prediction rather than the detrimental effects of integration difficulty, the best opportunity to identify the source of the N400 effect should arise in fMRI experiments that explicitly manipulate expectation instead of or in addition to congruity. Indeed, an fMRI contrast of non-anomalous unpredicted endings (*The pilot flies the kite*) to highly predictable endings (*The pilot flies the plane*) showed a main effect of expectancy in MTG (Baumgaertner et al., 2002), and another fMRI study showed parametric modulation of an MTG/IT area with expectancy of sentence ending (Dien et al., 2008). Another study created predictable contexts for congruent and anomalous sentences and showed both a robust N400 effect in ERP and significant fMRI activation in STG/STS (Kuperberg et al., 2003). These results suggest that the larger part of the classic N400 effect in semantic priming and anomaly is a consequence of facilitation of lexical access due to expectation, and not by anomaly per se (cf Hagoort et al., 2004).

Neither the semantic priming nor the semantic anomaly studies showed consistent effects in anterior temporal cortex (anterior STG or temporal pole). This follows from the model, which predicts anterior temporal effects when sentences are directly contrasted with words, but not when conditions differ only in the ease of accessing lexical representations—as the lexical access account of the N400 proposes for both the priming and anomaly manipulations. However, anterior temporal cortex is known to be susceptible to fMRI signal artifact (Devlin et al., 2000), particularly in the medial region that has been implicated in N400 generation by the intracranial studies discussed below (this medial region is distinct from the anterior STG region that shows sentence vs. word effects). Thus, we cannot rule out the possibility that this artifact masks real effects of semantic priming and anomaly in anterior temporal cortex.

MEG studies

MEG studies examining the N400 context effect have used use both equivalent current dipole models and distributed source models to estimate the source of the N400. Although achieving precise source estimates is difficult, MEG allows robust discrimination between activity generated in left vs. right hemispheres and anterior vs. posterior cortical areas over time, and can thus provide supporting evidence for the time-course of activity in areas identified with fMRI.

Studies of semantic priming and semantic anomaly using ECD source analysis uniformly report a response that localizes to left mid-posterior MTG/STS/STG and has an onset of around 250 ms in auditory presentation (Uusvuori, Parviainen, Inkinen, &

Salmelin, 2007; Helenius et al., 2002), and around 300-350 ms in visual presentation (Simos et al., 1997; Helenius et al., 1998; Helenius et al., 1999; Halgren et al., 2002; Pylkkänen and McElree, 2007; Service et al., 2007), and a peak latency of 410-450 ms[3]. Recent MEG studies have estimated the source of the semantic anomaly effect using a distributed source model based on the cortical surface (Halgren et al., 2002; Maess et al., 2006; Pylkkänen and McElree, 2007). In Halgren's study, early effects (between 250-500ms) were observed in the left planum temporale and left MTG/IT (Halgren et al., 2002). Additional areas—left anterior temporal and inferior frontal cortex and right orbital and anterior temporal cortex—were implicated in the later part of the anomaly response, which may reflect either the latter part of the N400 effect or the post-N400 positivity. However, other studies using distributed source models have found somewhat different results (Maess et al., 2006; Pylkkänen and McElree, 2007). Nevertheless, all methods of MEG source analysis have converged on the finding that the left mid-posterior temporal cortex is one source of the N400 effect. Furthermore, a recent study using simultaneous recordings of ERP and the event-related optical signal (EROS) corroborates the MEG source estimates: left-hemisphere responses to semantically anomalous sentence endings were observed in mid-posterior STS/MTG between 200-400 ms, with contributions from anterior temporal and inferior frontal areas only after 500 ms (Tse et al., 2007).

---

[3] A number of studies have described an MEG response known as the M350 which peaks between 300-400 ms and whose latency is sensitive to lexical factors such as frequency and phonotactic probability (Embick et al., 2001; Pylkkänen et al., 2002). The M350 refers to an evoked response to words, rather than to a difference between two evoked responses as in the case of the N400 context effect. Therefore, the N400-like context effect observed in MEG may involve differential activity during the same time-window as this evoked response, which would be consistent with our argument here that it reflects facilitation of lexical access, but the fact that the N400 effect is observed in the same time-window as the M350 does not necessarily constrain interpretations of the activity that gives rise to the M350, for the reasons discussed in Section 2.3.

Intracranial recordings

Further evidence on the source of scalp-recorded N400 effects derives from intracranial recordings collected from surface or depth electrodes on pre-operative epilepsy surgery candidates. The drawback of this method is that findings are limited to the areas that are plausible epileptogenic zones, often the anterior medial temporal lobe (AMTL) and inferior temporal areas.

Many studies recording from anterior medial temporal locations have found a 400-450 ms latency evoked response to visually presented words that shares many properties of the N400 response (Smith, Stapleton, and Halgren, 1986; Halgren, Baudena, Heit, Clarke, Marinkovic, & Chauvel, 1994a,b; Nobre & McCarthy, 1995; Elger et al., 1997). The amplitude of this AMTL response has been shown to correlate with the scalp-recorded N400 in semantic priming and semantic anomaly paradigms (Nobre et al., 1994; Nobre & McCarthy, 1995; McCarthy, Nobre, Bentin and Spencer, 1995). However, the restricted anatomical coverage of this technique has limited information about other regions. Intracranial recordings have demonstrated an N400 evoked response in areas including anterior inferior temporal cortex (Nobre, Allison, & McCarthy, 1994; Nobre & McCarthy, 1995) and prefrontal and orbitofrontal cortex and posterior lateral STG and MTG (Halgren et al., 1994a,b; Elger et al., 1997), and repetition priming effects have been demonstrated at MTG sites (Elger et al., 1997), but contextual manipulations have not been tested in these areas.

The evidence from intracranial recordings provides evidence that activity in AMTL correlates with the scalp-recorded N400 on several measures, raising the possibility that this area is an additional N400 generator. However, a non-generator region could also reflect N400 patterns if its function were partially contingent on the factors driving MTG activity. Indeed, intracranial recordings from AMTL demonstrate the opposite response to that of the scalp-recorded N400 to pronounceable nonwords (Nobre & McCarthy, 1995)—greater for words than pronounceable nonwords, in contrast to the scalp-recorded N400 that is typically larger for pronounceable nonwords than real words (Holcomb & Neville, 1990; Holcomb, 1993; Bentin et al., 1999). Unfortunately, fMRI data is inconclusive with respect to the contributions of this area to the N400 effect, as measurements in anterior temporal areas often suffer from artifact-induced signal loss or lack of coverage.

3.5 Conclusion

The goal of this chapter was to determine whether effects of context on lexical processing are partially due to top-down influences on activity in cortical areas subserving the storage and access of lexical information. I reviewed previous work that suggests that left posterior middle temporal cortex is the area most likely to support this function. The combined MEG and fMRI study that I presented in the first part of this chapter implicated different areas with MEG and fMRI: mid-posterior temporal cortex in MEG and inferior frontal cortex in fMRI. In the second part of this chapter, I argued that a comprehensive review of the previous literature leads to a better understanding of the source of the contextual effects that are reflected in N400 amplitude. My review and analysis of fMRI and MEG studies provides strong evidence that the N400 effect reflects reduced activity in posterior middle temporal cortex: this was the only area to show effects of semantic priming in fMRI across all the conditions that typically show an N400 effect, and MEG studies localize semantic priming and sentence context effects to the same place.

These data thus strongly suggest that a substantial part of the N400 context effect reflects a mechanism through which top-down information acts to reduce activity in lexical storage areas when the context is predictive of the current word. This finding has several important consequences. First, it provides strong evidence for a particular kind of predictive mechanism in language processing, and thus provides constraints on models of lexical processing in context. One could have imagined a model in which facilitative effects of supportive contexts are solely due to predictive structure building or predictive integration of predicted representations into higher-level structures without affecting the state of the stored lexical representation; however, such a model is not supported by these data. Second, this finding suggests that we may be able to use the N400 effect as an index of top-down processing, which would be a huge methodological advance. However, we first need to determine whether the N400 effect reflects the summation of multiple functional components, and if so, find a way of dissociating them.

It is important to note that these data do not rule out the possibility that other kinds of predictive mechanisms also facilitate processing in the same situations, nor do they rule out the possibility that the N400 context effect reflects more than one component process, and in turn, more than one generator. In addition to the posterior middle temporal effects, a subset of studies also show effects in inferior frontal regions, possibly reflecting differences in ease of selection or retrieval or predictively facilitated integration of the current word with syntactic, thematic, or discourse structures. These effects may be reflected in a subsequent late positivity, as I will discuss in more detail in the next chapter, but they may also contribute to the N400 effect. The anterior temporal cortex may also contribute to the scalp-recorded N400 but fail to be represented in fMRI data for method-specific reasons. Such issues all merit further investigation.

In the next chapter, I use the results of the meta-analysis as a foundation for a partial neuroanatomical model of language comprehension. In particular I focus on the role of left anterior IFG, which I argue to be the region mediating top-down and predictive activation of stored lexical-semantic representations in posterior MTG.

4 Neuroanatomy of processing words in context

4.1 Introduction

In this chapter, I will sketch out some of the pieces of a neuroanatomical model of processing words in context that incorporates a mechanism for predictive pre-activation. I should say from the outset that I don't think we should expect that all the functional operations required for processing words in sentences are associated with a particular area of cortex, or at least an area that is localizable by current methods. If not, it doesn't make sense to try to find a brain area for each of the operations that we know to be required on theoretical grounds. It's also not necessarily the case that, if a brain region's level of activity is correlated with the degree to which a given computation is required, that this is the brain region that actually does the computation of interest rather than playing some kind of supporting role. I think what we can realistically do under these circumstances is to take the areas of cortex that have been empirically demonstrated to have a functionally selective profile across a number of neuroimaging methods and see if they match up with operations that are required by our processing theory. To the extent that we find some regions that do correlate with some of these operations, we can at the least begin to use them as a new dependent measure for testing theoretical questions that involve the use of these operations, and we can at the same time try to develop more subtle tests to discover whether the regions of interest actually play a causal role.

I will assume a very underspecified model of the operations required for lexical processing in context. For the bottom-up route, I assume that auditory or visual input goes through early processing that leads to activation of orthographic or phonemic representations, which in turn leads to activation of lexical representations that are bindings of phonological, syntactic, and conceptual information. I assume that one among the many activated lexical representations must be selected for further computations, and that the appropriate information from this selected representation is added to syntactic and compositional semantic representations of the previous sentence fragment, and to some kind of representation of the situation or discourse, all of which are being built incrementally as each word is encountered. For the top-down route, I assume that information from these higher-level representations must be used as input into some process that uses prior knowledge to determine which lexical, conceptual, or syntactic representation is likely to be coming up. This prediction may affect a number of the bottom-up steps for processing the next input; based on the results in the past two chapters, I will assume that one of the steps affected is the activation of lexical information.

I identified three main cortical areas that are selectively associated with certain aspects of language processing across the literature: left inferior frontal cortex, posterior temporal cortex, and anterior temporal cortex. I also included left angular gyrus, an inferior parietal region that has recently been implicated in semantic processing. Based on a survey of prior literature and a meta-analysis of fMRI studies of semantic priming and sentence comprehension in the previous chapter, I will propose the following model. Different parts of a temporo-parietal network are responsible for (i) storing long-term representations of words (posterior temporal cortex) and (ii) constructing and maintaining temporary representations of the larger syntactic and semantic structure (anterior temporal and inferior parietal cortex) (Damasio, 1991; Damasio & Damasio, 1994; Hickok & Poeppel, 2004, 2007; Martin, 2007). Inferior frontal cortex controls the flow of information between these systems by (i) guiding the activation of stored lexico-semantic information based on the context (anterior IFG) and (ii) managing the selection of candidate representations with which to update the context (posterior IFG) (Badre & Wagner, 2007; Thompson-Schill et al., 2005; Gabrieli et al., 1998).

In the previous chapter I reviewed the evidence that lexical representations are stored in posterior middle temporal cortex, and I showed that the availability of context-based lexical prediction impacts activity in this area. Since this thesis is centered on the role of predictive mechanisms in language comprehension, I will begin this chapter by presenting extensive evidence for the hypothesis that anterior IFG supports top-down activation of these stored representations. I will discuss how this capability could be used to implement predictive pre-activation of lexical and conceptual representations during language comprehension, and how this hypothesis can explain the pattern of activity observed in this area across the priming and sentence context experiments reviewed in the last chapter. Next I will discuss evidence that posterior IFG supports selection between competing activated representations. Finally, I will briefly discuss the possible role of anterior lateral temporal cortex and angular gyrus in combinatorial operations.

4.2 Retrieval and anterior inferior frontal cortex

Anterior inferior frontal cortex and semantic processing

For a number of years, it has been observed that ‘semantic’ manipulations in fMRI are associated with increased activity in left prefrontal cortex (see Bookheimer, 2002, for review). A sampling of the kinds of tasks and contrasts that used to elicit this response is presented in Table 5.

	experimental task	**control task**
Petersen et al. 1988 (PET)	given a noun, generate an associated verb	read noun out loud
Kapur et al. 1994	given a word, decide whether it is living or non-living	decide whether word contains a particular letter
Demb et al. 1995 Ex. 1	given a word, decide whether it is abstract or concrete	decide whether word is in uppercase or lowercase
Demb et al. 1995 Ex. 2	given a word, decide whether it is abstract or concrete	decide whether first and last letter of word are in alphabetic order or not
Vandenberghe et al. 1996 (PET)	decide which of two words or pictures on bottom of screen are more like word or picture on top in meaning	decide which of two words or pictures on bottom of screen are closer in size to stim-identical word or picture on top
Gabrieli et al. 1996	given a word, decide whether it is abstract or concrete	decide whether word is in uppercase or lowercase
Spitzer et al. 1996	given a pair of words, decide if they are related in meaning	decide whether rows of asterisks are same or different color

Table 5. Early neuroimaging studies that found ‘semantic’ effects in IFG, and the tasks that were contrasted to elicit the effects. Studies used fMRI unless otherwise noted.

Most of these early studies contrasted semantic tasks with low-level perceptual baselines, and showed differential activity throughout left IFG. A number of subsequent studies have contrasted phonological and semantic tasks and have consistently shown a pattern in which left anterior/ventral IFG (particularly pars orbitalis – BA 47) is selectively activated for semantic tasks, while more posterior/dorsal areas of IFG (BA 44/45) tend to be activated for both or selectively for phonological tasks. Figure 19 from Badre and Wagner (2007) illustrates these different sub-regions of inferior frontal cortex, and Table 6 lists some of the particular manipulations used to dissociate these regions.

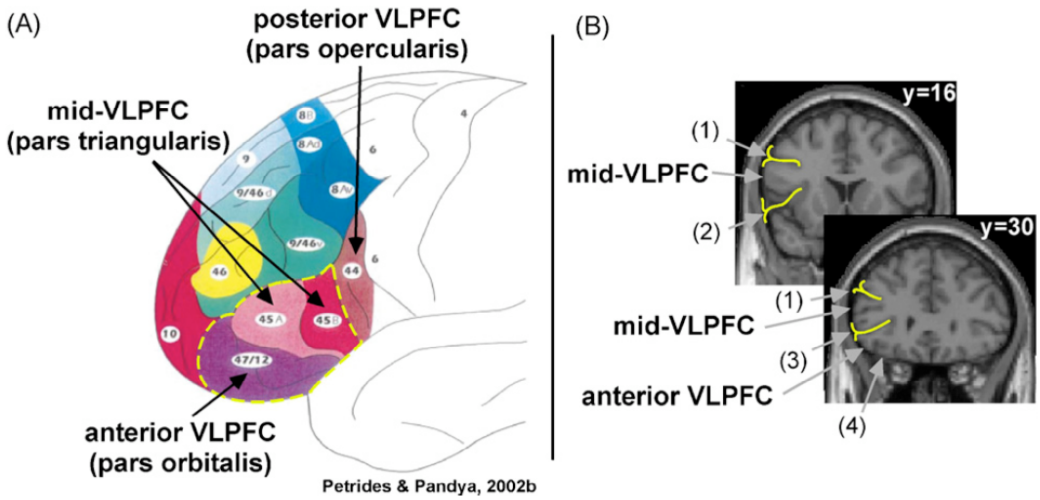


Figure 19. Anatomical divisions of inferior frontal cortex, also known as ventrolateral prefrontal cortex (VLPFC) from Badre & Wagner, 2007, adapted from Petrides & Pandya, 2002.

	semantic task	**phonological task**
Poldrack et al. 1999	given a word, decide whether it is concrete or abstract	given a word, count the number of syllables
Roskies et al. 2001	given a pair of words, decide if they are synonyms	decide whether pair of words rhyme
McDermott et al. 2001	attend to relation between words in list of semantically related words	attend to relation between words in list of rhyming words
Gough et al. 2005 (TMS)	given a pair of words, decide if they are synonyms	given a pair of words, decide if they are homophones
Mechelli et al. 2007	name words in pairs that happen to be semantically related	name words in pairs that happen to be phonologically related

Table 6. Neuroimaging studies that reported dissociation between anterior and posterior IFG by contrasting semantic and phonological tasks, and the tasks that were used. Studies used fMRI unless otherwise noted.

Other findings implicate anterior IFG in particular in tasks that involve processing of lexical semantics. Dapretto and Bookheimer (1999) showed that when judgment of meaning identity in sentences depended on lexical semantics rather than syntactic transformations, significantly more activity was observed in anterior IFG. fMRI studies contrasting word and pseudoword processing sometimes find more activity selectively in anterior IFG for words relative to nonwords (Binder et al., 2003; Orfanidou et al. 2006) although not always (Price et al., 1996; Mechelli et al., 2003).

aIFG and semantic retrieval

The fact that all of these contrasts involved access of stored semantic information led to the early proposal that aIFG supports some aspect of retrieving information from semantic memory[4] (Buckner et al., 1995; Demb et al., 1995). Memory retrieval, as defined in a recent review, refers to the 'processes that bring information from the past back into the cognitive focus'. An obvious property of retrieval processes is that they require as input at least one property of the information desired—a 'cue'. This cue may be as rich as the sensory context in which the information was acquired, or may be as minimal as a time or sequence index (retrieve the stimulus I saw two trials earlier). In the verb generation task, the cue is the noun presented combined with the context of the task, which is to retrieve a verb related to the presented noun. In the synonym task, the cue is the two words combined with the context of the task, which is to retrieve the basic meaning of each word.

These early proposals did not specify the role that aIFG played in retrieval of semantic information. If aIFG were involved in any kind of access of semantic memory, the prediction would be that aIFG activity should be observed in response to any stimuli like words that automatically trigger access of semantic information, unless the task specifically discouraged processing words for meaning. However, in a very influential paper, Thompson-Schill and colleagues (1997) pointed out that some tasks that would

---

[4] 'Semantic' here should be interpreted according to the distinction in the memory literature between semantic, or declarative knowledge and other kinds of stored information such as episodic or procedural memory.

seem to involve this kind of automatic semantic access had failed to demonstrate aIFG activity. In their own study, Thompson-Schill and colleagues tested a related prediction, that experimental conditions that required access of semantic information for four compared to two words should lead to greater IFG activity, and also did not find a significant effect.

Thompson-Schill and colleagues (1997) argued that left IFG subserves selection between competing representations rather than retrieval, but subsequent work has shown that such selection effects are limited to mid-IFG (BA 45), as I will discuss in more detail in the next section. However, Thompson-Schill et al.'s findings led Wagner and colleagues (2001) to propose a slightly more constrained role for aIFG. They suggest that left aIFG (what they refer to as LIPC) supports not all retrieval of semantic information from memory but specifically what they call *controlled* retrieval:

> I propose that LIPC (lateral inferior prefrontal cortex) contributes to controlled semantic retrieval. That is, LIPC mechanisms may guide the recovery of semantic knowledge under situations where pre-experimental associations or prepotent responses do not support the recovery of task-relevant knowledge through more automatic mechanisms. When a strong association exists between two elements, be they two stimuli or a stimulus and a response, the presentation of the first element may yield sufficient activation of the second element such that this associated representation may be accessed relatively automatically. That is, the second element may be recovered even in the absence of top-down facilitation or bias. Importantly, considerable evidence suggests that prefrontal regions are particularly important for cognition and behavior under conditions where strong stimulus-stimulus or stimulus-response associations are absent…The increased role of prefrontal cortex when associations are weak may reflect the greater need for top-down bias signals to guide controlled access to or retrieval of the associate when presented the first element (Wagner et al., 2001; p. 330).

Wagner et al. (2001) argued that in Thompson-Schill et al.'s (1997) materials, strong associative relationships between the probe word and target could have been detected automatically on both two- and four-candidate trials, mitigating the need for any significant amount of controlled semantic retrieval. Wagner and colleagues found that when materials with a weaker association were used, left aIFG showed a significant effect of retrieval demand, with greater activity in the 4-choice than the 2-choice condition; they also found greater activity in left aIFG in the 2-choice weak-association condition than the 4-choice strong association condition, bolstering their argument that a different mechanism (automatic vs. controlled retrieval) was engaged in the strong-association and weak-association conditions. The same pattern of aIFG activity was subsequently replicated by Badre et al. (2005) with a similar manipulation.

Therefore, taken together, the findings of Thompson-Schill et al. (1997) and Wagner et al. (2001) support the claim that left aIFG is involved in retrieval of stored semantic information when this information is not automatically made available through the bottom-up activation of stored representations. This view provides a natural anatomical basis for the classic behavioral findings of distinct patterns of 'automatic' and 'strategic' lexical semantic retrieval at short and long SOAs. At short SOAs, the only semantic information that can be retrieved on the basis of the prime is information that has strong connections to the prime within the level of long-term memory representations in posterior middle temporal cortex. At long SOAs, semantic information can be retrieved based on knowledge of the task and knowledge of within-experiment regularities (e.g., words for birds predict words for buildings) that may be represented outside of the semantic network, and which informs top-down retrieval through aIFG.

Top-down connectivity between aIFG and MTG

In Chapter 3, I reviewed evidence that lexical and conceptual information is stored in left posterior temporal cortex. Therefore, it would seem that for aIFG to mediate top-down retrieval of semantic information, it must have connections with posterior temporal cortex through which it can bias activity over particular representations. There is at least some evidence for such connections. Petrides and Pandya (2002) showed using retrograde tracing that the area corresponding cytoarchitecturally to BA 47 in monkeys has strong inputs from posterior inferotemporal areas. More recently, Croxson and colleagues (2005) used diffusion-weighted imaging tractography to examine human and monkey PFC connectivity. They found that a more ventral area of IFG had the strongest connections with posterior temporal regions while a more dorsal and posterior area of IFG had the strongest connections with parietal regions. Although all of this work is still very preliminary, it suggests at least that the anatomical connections required by the targeted semantic retrieval hypothesis are likely to be available.

In terms of functional connectivity, Bokde and colleagues (2001) found that the correlated activity between left aIFG and posterior temporal areas was more semantic-specific (stronger for reading words than pseudowords) than for dorsal IFG. Similarly, Saur et al. (2008) used a diffusion tensor imaging analysis during repetition and comprehension and found evidence for a functional pathway between left aIFG and posterior middle temporal cortex during the comprehension task, in contrast to a pathway between posterior IFG and superior temporal cortex during the repetition task. Co-dependence between activity in aIFG and in posterior temporal cortex is also observed across a number of manipulations. Wagner et al. (2001) and Badre et al. (2005) both found that the same contrast in retrieval demands that elicited differential ventral IFG activity also elicited differential activity in left middle temporal cortex. In a study of episodic memory, Dobbins and Wagner (2005) found that both left middle temporal cortex and ventral IFG showed more activity for recollection of conceptual details than perceptual details, and that the magnitude of this difference in the two regions was significantly correlated across subjects. Similarly, the Gold et al. (2006) semantic priming study discussed above found that activity in aIFG and middle temporal cortex tracked each other across manipulations of relatedness relative to a neutral condition. Staresina et al. (2009) found that increased activity in left aIFG correlated with increased activity in posterior inferior temporal cortex. Finally, a number of studies show reductions in both left aIFG and middle temporal cortex when items are repeated in semantic tasks (e.g. Raichle et al., 1994; Buckner et al., 2000). In fact, outside of the sentence context studies I reviewed in Chapter 3, there are few fMRI studies that show effects in aIFG and not in posterior temporal cortex. On the other hand, a number of studies (e.g. Badre et al. 2005, the short SOA semantic priming experiments reviewed in Chapter 4) demonstrate effects in posterior middle temporal cortex and not in aIFG under automatic retrieval conditions, which would be expected if posterior temporal activity can be modulated by both automatic activation from the input and top-down activation from aIFG.

One undesirable consequence of this account of aIFG is that it seems to require that aIFG in some sense recreates the entire semantic network, so that it can pre-activate any particular representation when the context supports it. So far I do not have a good response to this objection, but there is at least some empirical evidence that prefrontal cortex has the necessary properties to play this role. In their seminal review of prefrontal cortical function, Miller and Cohen (2003) summarize converging evidence that prefrontal neurons (1) receive inputs from all over cortex, (2) can represent item and category information that is also represented (although perhaps with greater specificity) in posterior temporal cortex, (3) can represent specific associations between items, contexts, and responses, (4) can affect activity of neurons in posterior temporal cortex through top-down connections. Although most of this evidence comes from experiments on visual object recognition in primates and involves other areas of prefrontal cortex besides aIFG, it shows that a hypothesis of top-down activation of semantic representations that requires aIFG to have these properties is not unreasonable.

aIFG and prediction

I have reviewed evidence that left aIFG mediates targeted semantic retrieval, and that it does this by exerting a top-down bias on activity on the stored semantic representations themselves in posterior middle temporal cortex. In the work reviewed, aIFG has mainly been argued to mediate its effects after presentation of the critical stimulus, by retrieving semantic information associated with the basic lexical or conceptual representation activated by the bottom-up input, often based on relevance to the task at hand. However, it should be obvious by now that this kind of mechanism could equally support predictive pre-activation of stored representations before the bottom-up input is presented. Badre and Wagner (2007) suggest that targeted semantic retrieval is necessary when automatic retrieval processes are not sufficient for bringing the relevant information into focus, and the pre-stimulus time-window could be seen as an extreme case, in which there is not yet any sensory input available to initiate automatic retrieval.

In language comprehension, predictions often need to be maintained across intervening material: for lexical prediction cases like those in N400 sentence studies, adjectival modifiers could intervene between a verb and a predicted noun; for syntactic dependency cases as I will discuss in Chapter 6, the predictions could need to be maintained across several clauses. It is a well-known property of prefrontal cortex in general that it can maintain representations across a delay (Miller & Cohen, 2003), although this has not been demonstrated for aIFG in particular.

I propose, therefore, that during sentence comprehension left aIFG is involved in predictively retrieving stored lexical and conceptual representations on the basis of the prior lexical, sentential, and discourse context. This pre-activation facilitates automatic retrieval of the lexical and conceptual representations when the critical input is subsequently presented. An advantage of this proposal is that it does not need to posit a special area to mediate predictions, which would seem unlikely anyway. However, because aIFG is involved in targeted semantic retrieval both before and after bottom-up input is encountered, aIFG effects do not necessarily reflect predictive effects. In fact, although I have just presented several reasons to believe in theory that aIFG mediates predictive pre-activation during processing, in the next section I will argue that most of the aIFG effects observed thus far in language comprehension may reflect non-predictive targeted semantic retrieval.

aIFG effects in fMRI of semantic priming

In the meta-analysis presented in Chapter 4, I showed that a number of semantic priming experiments found a priming-related reduction in left aIFG activity, but that such an effect was only observed when the SOA between prime and target was fairly long (> 600 ms). To illustrate these effects, Figure 20 presents representative left aIFG contrasts (normalized BOLD signal relative to visual fixation baseline) from the study by Gold and colleagues (2006).

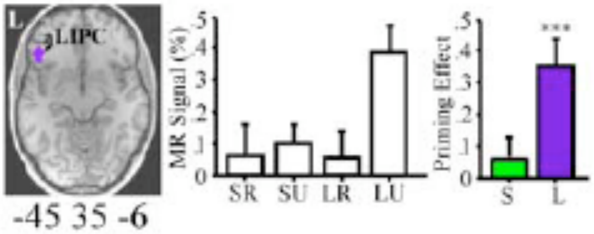


Figure 20. Results from Gold et al., 2006, Experiment 1, which examined the response to semantically related and unrelated word pairs. Bar charts show mean BOLD percentage signal in left aIFG relative to a visual fixation baseline for each experimental condition. SR = short SOA, related, SU = short SOA, unrelated, LR = long SOA, related, and LU = long SOA, unrelated. The two colored bars illustrate the size of the priming effect (unrelated-related) in short (S) and long (L) SOA conditions.

The results of Gold et al. (2006) show that activity in left aIFG was not significantly greater than baseline in either the short-SOA condition or the long-SOA condition. This argues against an account of these data in which the left aIFG activity reflects a mechanism associated with basic lexical access that is facilitated in predictive contexts, because that would predict significant activity in both short- and long-SOA unrelated conditions, as both targets were unpredictable based on context.

I suggest that this data is best accounted for by assuming that left aIFG is engaged when there is enough time between prime and target to use the prime as a predictive top-down retrieval cue, in the long SOA conditions. In the related long-SOA condition, the target will either be accessed predictively through targeted semantic retrieval before the target is presented, or quickly and automatically based on strong associative links after the target is presented; either way, no effort will be necessary to retrieve the target or the relationship between prime and target after the target is presented. In contrast, in the unrelated long-SOA condition, the target that is presented will not match the retrieval cues. Therefore, the retrieval process continues longer, perhaps because it continues for some pre-specified amount of time until it finds a match, or perhaps in a vain effort to retrieve some stored relationship between the retrieval cue and the target.

Our account suggests that the pattern of left anterior IFG activity observed in priming experiments does not directly reflect predictive processing, but rather is due to a post-access search for the relationship between prime and target. Does this mean that the left anterior IFG effect is just a 'task effect'? On the contrary, I would argue that the semantic priming paradigm happens to tap into a mechanism that is central to successful sentence comprehension, the attempt to retrieve stored information about the relationship between two words presented in succession.

aIFG effects in fMRI of sentence processing

A number of the sentence-level fMRI studies reviewed in Chapter 3 reported effects of contextual fit in left aIFG. These studies contrasted comprehension of normal sentences with sentences containing words that were semantically or pragmatically incongruent with the preceding sentence context, and reported more activity in aIFG for sentences containing incongruities. All but two studies reported an effect somewhere in left aIFG[5], but effects were reported in all three sub-parts of aIFG across different studies. Only 9 of these studies specifically reported effects in BA 47, the most anterior and ventral sub-part of IFG, and the locations of these effects within IFG were not obviously related to particular modality or task parameters. However, as these studies were mostly not designed to distinguish between different areas, many simply reported the location of the center of a cluster of significant IFG activity rather than delineating the full extent of the cluster[6]. Therefore, it may well be the case that incongruity routinely gives rise to

---

[5] No obvious design parameter distinguished these two experiments from the others; in particular, Friederici et al. (2003) used a design very similar to other studies that did report an LIFG effect (Rüschemeyer et al., 2005).

[6] Another problem is substantial variability in the location of inferior frontal structures across participants (Amunts et al., 1999; Fedorenko & Kanwisher, submitted), although it's not clear that this should be more

differential activity in both anterior and posterior IFG, as was observed by a recent study that isolated BA 47 and BA 45 in separate ROIs (Menenti, Petersson, Scheeringa, & Hagoort, 2009).

Computing the appropriate interpretation of a sentence beyond basic predicate-argument relationships requires accessing stored knowledge about the relationship between words, regardless of whether processing is predictive or not. The hypothesis that left aIFG is involved in retrieval from semantic memory therefore predicts that aIFG should normally be active during sentence comprehension, regardless of the task assigned. In support of this, several studies have reported left aIFG activity during sentence comprehension relative to a low-level baseline when no task beyond comprehension is given (Crinion et al., 2003; Lindberg & Scheef, 2007), although several others do not (Giraud et al., 2004; Spitsyna et al., 2006). A number of the semantic anomaly studies included in the meta-analysis also report significant increases in activity in this region for all sentences relative to a low-level baseline (e.g. Kuperberg et al., 2003; Cardillo et al., 2004), as did we in the experiment presented in Chapter 3.

If some amount of controlled semantic retrieval can be expected during the processing of most content words in order to activate the appropriate semantic information given the context, then experimental effects in this region for processing words in sentences can go in two ways—activity could be reduced or increased relative to the baseline. When the next word of the sentence satisfies the prediction, targeted retrieval should be minimized—the salient information needed to relate the current word and the previous context must have already been retrieved for the word to be predicted.

---

of a problem in sentence experiments than in semantic judgment experiments that do succeed in functionally isolating anterior and posterior IFG with group analyses across an average brain.

When the next word in a sentence results in an interpretation that does not fit with world knowledge, as in incongruous materials, an obvious repair strategy is to try harder to retrieve some stored relationship between the previous context and the current word that will make the combination fit with world knowledge.

Most fMRI studies have not explicitly crossed predictability and congruity, so it is hard to determine whether one or both factors are the cause of the differences in left aIFG activity observed in the fMRI studies of semantic anomaly reviewed in Chapter 4. One experiment by Baumgaertner and colleagues (Baumgaertner, Weiller, & Büchel, 2001) contrasted the response to expected (*The pilot flies the plane*), unexpected (*The pilot flies the kite*) and incongruous (*The pilot flies the book*) sentence endings (a fourth condition was pseudoword endings). The authors reported a significant difference between anomalous and unexpected endings in left aIFG, but all other contrasts were non-significant in this region. However, the relative magnitude of aIFG signal change across the conditions was exactly as I would predict, as illustrated in Figure 21 (the hemodynamic response function is modeled against a background of visual fixation between trials). The anomalous condition showed the greatest activity, while the unexpected condition (which I take to be the baseline, normal case) showed less activity, but appears to be substantially above zero. If the expected condition requires no additional retrieval of semantic information, it could feasibly demonstrate no significant activity above baseline, as is observed here. This data is suggestive, though preliminary; clearly the impact of predictability on left aIFG response needs to be examined more systematically to test this hypothesis.

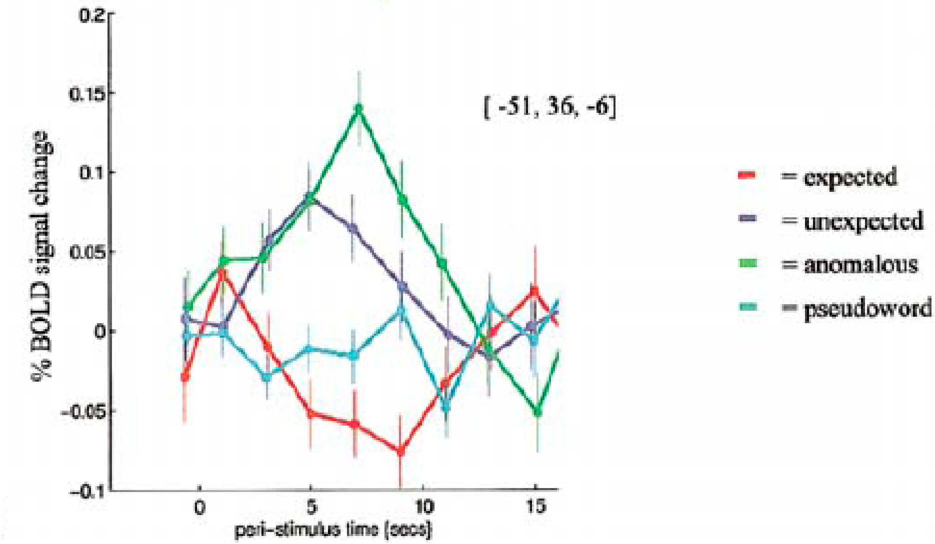


Figure 21. % BOLD signal change averaged across a left inferior frontal ROI centered on aIFG from Baumgaertner et al. (2002), modeled against a baseline of visual fixation. The conditions compared were sentences of the type *The pilot flies the plane* (expected), *The pilot flies the kite* (unexpected), *The pilot flies the book* (anomalous) and *The pilot flies the lurge* (pronounceable pseudoword).

4.3 Selection and posterior inferior frontal cortex

While I have argued that anterior IFG (BA 47) supports targeted semantic retrieval, there is increasing evidence from a number of literatures that a more posterior and dorsal area of left IFG (BA 45) supports domain-general resolution of competition among activated candidate representations (Thompson-Schill et al., 1997; 1998; Badre et al., 2005; Novick, Trueswell, & Thompson-Schill, 2005; Badre & Wagner, 2007; Kuhl & Wagner, 2009). Because it has been suggested that the most posterior part of IFG (BA 44) may be involved in phonological perception and production (e.g. Hickok & Poeppel, 2007; Badre & Wagner, 2007), I refer to the BA 45 area thought to be involved in selection as mid-IFG (mIFG).

First, memory studies of proactive interference frequently associate this region with interference resolution (Jonides & Nee, 2006; Jonides et al., 2007; Kuhl & Wagner, 2009). Proactive interference refers to the idea that items in memory compete to be in the focus of attention, and that competition from items from the past may interfere with encoding and retrieval of new items, especially when old items and new items are similar. A number of PET and fMRI studies have found increased activation in this area when interference is increased. For example, in one classic memory task, participants first have to learn a set of pairings (DOG-BOXER) and then in the second part of the task have to either learn completely new pairings (CAR-BOOKSHELF) or rearrangements of the old pairings (DOG-DALMATIAN). Learning the new associations for previously studied items is thought to increase interference, and shows increased activity in left mIFG relative to learning pairings with new items (Dolan & Fletcher, 1997; Henson, Shallice, Josephs, & Dolan, 2002). A number of studies show this increase in mIFG activity occurring during the response period, when contrasting conditions under which a probe matched a recent non-target, thus requiring interference resolution (e.g. Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998; D'Esposito, Postle, Jonides, Smith, & Lease, 1999) and when comparing memory retrieval across more or less intervening items (Nee & Jonides, 2008; Öztekin et al., 2008).

Second, researchers using semantic judgment tasks that require semantic retrieval find greater activity in left mIFG when the task requires selection (Thompson-Schill et al., 1997; Badre et al., 2005). For example, judging which words are most similar along one particular feature dimension requires selective attention to one dimension of similarity in a way that judging which words are most 'related' across all dimensions does not. Left mIFG is also more highly activated in verb generation tasks for stimuli which are likely to elicit multiple different responses relative to stimuli for which one response is preponderant (e.g., Crescentini, Shallice, & Macaluso, 2009), and patients with left mIFG damage have greater deficits in verb generation for stimuli with high selection demands (Thompson-Schill et al., 1998).

Third, recent fMRI studies of lexical ambiguity have consistently found greater activity in left mIFG for ambiguous words, for which one of several possible meanings must be selected, than for unambiguous controls (Rodd et al., 2005; Mason & Just, 2007; Zempleni, Renken, Hoeks, Hoogduin, & Stowe, 2007; Bilenko et al., 2008; Bedny, McGill, & Thompson-Schill, 2008). Patients with left mIFG damage also show an abnormal pattern of reaction time effects in contextual manipulations of homonymy and polysemy (Metzler, 2001; Bedny, Hulbert, & Thompson-Schill, 2007). Other recent studies find increased activity in the same area for sentences containing syntactic ambiguity (Mason, Just, Keller, & Carpenter, 2003; January, Trueswell, & Thompson-Schill, 2009). January and colleagues further show within subjects that this same left mIFG area shows increased activation for both syntactic ambiguity and for response ambiguity due to competing conflicting information in the Stroop task.

While much research links left mIFG with contexts in which candidate representations or responses compete, the underlying mechanism is still unresolved: competition, selection, inhibition of alternatives? Recent work by Grindrod and colleagues provides some evidence that it is not just competition but the resolution of competition that this area is sensitive to; they find that ambiguous words only show increased mIFG activity when the context resolves their meaning (Grindrod, Bilenko, Myers, & Blumstein, 2008). However, it remains to be understood how left mIFG is involved in this resolution of competition (see Nee & Jonides, 2006, for a number of possibilities).

I propose that the left mIFG effects observed in the semantic priming and sentence context experiments reviewed in Chapter 3 can be explained as reflecting resolution of competition between the predicted and actual input. In supportive contexts, the predicted and actual inputs are the same or similar, and there is little conflict. However, in anomalous sentences, the context is likely to predict input that is very different from the actual anomalous continuation. Similarly, in a priming experiment where many of the pairs are composed of related words, the first word is likely to elicit a prediction for the second word that will conflict with the actual input on unrelated trials. This can account for why left mIFG frequently shows more activation for anomalous relative to congruent sentence endings and for unrelated relative to related word pairs. Left mIFG effects are only observed in priming experiments for long SOAs because only then is there enough time for a potentially conflicting prediction to be formed. This hypothesis also explains why Gold et al. (2006) found an increase in left mIFG for unrelated relative to neutral pairs; as no prediction is formed in the neutral pairs, there is no conflict.

Left mIFG (BA 45) seems to overlap with what has classically been referred to as 'Broca's area' (usually BA 44 & 45), traditionally implicated in speech production and syntactic processing. It may be that these regions are actually anatomically distinct; as I mention above, BA 44 is likely to be involved in phonological processing and perhaps other functions, and it is possible that BA 45 may be broken down further into functionally distinct sub-areas. However, it may also be that some of the manipulations designed to differentially induce syntactic mechanisms may have inadvertently manipulated selection demands (Novick et al., 2005; January et al., 2009). More research is needed to distinguish these possibilities.

4.4 Beyond the word: lateral anterior temporal cortex

Left lateral anterior temporal cortex (ATC) has been implicated by many functional imaging studies of language processing in recent years. In this section I will discuss several important findings that emerge from this work. First, although the nature of the computations supported by left ATC are still unclear, the evidence strongly suggests that this region plays some role in the processing of structured representations above the single word level. Second, left ATC may support processing of structured representations across domains, although different sub-areas of ATC may be devoted to particular kinds of information. Here, ATC will refer to lateral anterior STG, STS, and MTG (anterior BA 22/21) and the lateral temporal pole (BA 38).

A large number of PET and fMRI studies have demonstrated significantly greater signal in bilateral ATC when sentence comprehension is contrasted with low-level baselines like consonant strings and reversed speech (Bavelier et al., 1997; Crinion, Lambon-Ralph, Warburton, Howard & Wise, 2003; Noppeney & Price, 2004; Spitsyna et al. 2006), as well as in contrasts between sentences and lexical-level baselines like reading or listening to word lists (Mazoyer et al., 1993; Stowe et al., 1998; Friederici, Meyer, & von Cramon, 2000; Humphries, Binder, Medler, & Liebenthal, 2006).

Although these studies show bilateral ATC effects, other evidence suggests that effects in ATC have different functional correlates in the two hemispheres, and in particular that right ATC preferentially supports processing of prosodic information. While left ATC differentiates intelligible from unintelligible speech, right ATC responds preferentially to speech with a normal intonation and perceived pitch, whether or not it is intelligible (Scott, Blank, Rosen, & Wise, 2000; Spitsyna et al., 2006). Right ATC has also been implicated in music processing (Griffiths et al., 1998) although at least one sub-part responds more to sentences than melodies (Rogalsky, Saberi, & Hickok, 2009). On the other hand, one of the few studies to contrast sentences and word lists using visual stimuli found a left ATC effect only (Vandenberghe, Nobre, & Price, 2002), and other studies have reported selective effects of sentence structure in left ATC (Humphries, Love, Swinney, & Hickok, 2005; Humphries, Binder, Medler, & Liebenthal, 2006; Rogalsky & Hickok, 2008). Therefore, while ATC in both hemispheres may be involved in processing supralexical structures, right ATC may be devoted to prosodic structure and left ATC to some level of structure that depends on intelligibility--syntactic, semantic, thematic, or discourse (although see Ferstl et al., 2008, for arguments that right ATC also plays a role in discourse processing).

Selective damage to left anterior superior temporal cortex has been associated with significant comprehension impairment for most sentence types more complex than simple declaratives (Dronkers et al., 2004; although cf Kho et al., 2007). Left ATC shows reduced activation in contexts that support syntactic priming (Noppeney & Price, 2004) and Brennan and colleagues (submitted) show that activity in left ATC correlates with a measure of syntactic complexity. Furthermore, left ATC shows a preference for sentences over word lists even when the stimuli are composed of pseudowords lacking in lexical semantic content (Friederici et al, 2000; Humphries et al, 2006). However, most of these findings do not definitively show that left ATC is involved in processing or representing syntactic structure rather than other levels of sentence- or discourse-level structure since complexity at one of these levels is so often associated with complexity at the others, and some form of thematic and discourse-level processing could still be done with pseudowords lacking lexical semantic context. All that we can conclude at this point is that left ATC plays a role in processing linguistic representations above the word-level

Anterior temporal regions have sometimes also been implicated in lexical-level processing, mainly on the basis of semantic dementia (SD), a disorder characterized by bilateral anterior temporal atrophy. However, since recent work suggests that the atrophy extends to other parts of the temporal lobe as well (Gorno-Tempini et al., 2004; Mummery et al., 2000; Williams, Nestor & Hodges, 2005; Noppeney et al., 2007), and since SD patients show reduced activation of posterior IT areas during semantic tasks (Mummery, Patterson et al., 1999), I think it more likely that the lexical semantic deficits in SD are largely due to reduced functionality of posterior temporal areas, whether due to atrophy per se or to loss of connectivity.

It is important to note that the lateral ATC region where the fMRI and PET effects of sentence structure have been reported is quite different from the medial anterior temporal region from which intracranial recordings displaying an N400 effect have famously been reported (McCarthy et al., 1995; Nobre et al., 1995). For illustration, I show the authors' MRI tracing of one of their subjects in Figure 22. Electrodes 3 and 4 in the figure, on the ventral anterior temporal surface referred to as anterior medial temporal lobe (AMTL) in this study, showed an N400 effect. The neuroimaging studies reported above showed effects on the lateral surface of STG and MTG (labeled on the right hemisphere of this figure). These areas may or may not share the same function; since fMRI studies suffer from significant signal loss on the anterior ventral surface, it is difficult to know.

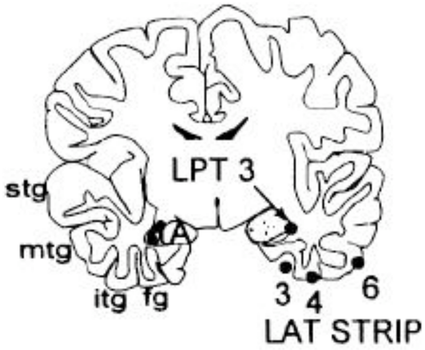


Figure 22. MRI tracing from Nobre and McCarthy (1995) showing the position of intracranial electrodes from which ERPs showing similar response properties to the N400 were recorded in one subject (indicated by the dots on the right side of the figure).

4.5 Beyond atomic semantic representations: angular gyrus

The angular gyrus (BA 39) is not known as one of the ‘classic’ language areas. However, evidence is accruing that this area is somehow involved in the processing or representation of the combination of atomic representations from semantic memory. Lesions here are associated with difficulty in processing complex sentences (Dronkers et al., 2004), and imaging studies demonstrate greater AG activity in response to words than nonwords (Binder et al., 2003, Binder, Medler, Desai, Conant, & Liebenthal, 2005; Ischebeck et al., 2004; Rissman, Eliassen, & Blumstein, 2003; Mechelli, Gorno-Tempini, & Price, 2003) and in response to semantically congruent sentences than semantically related word lists or syntactically well-formed but meaningless sentences (Humphries et al., 2007). Increased AG activity is also found for mismatches between visually presented words/pictures and environmental sounds, suggesting that it may be involved in integration of semantic information from both linguistic and non-linguistic sources (Noppeney et al., 2008).

By itself, this handful of studies is perhaps rather slim evidence for AG playing a major role in semantic combination. However, Binder and colleagues (1999, 2009) have pointed out that AG is also part of the ‘default network’; a network of areas that has been identified as being active when participants are ‘resting’ in the MRI scanner (as all fMRI responses are defined relative to a baseline, these have been identified as areas that show a reduction in activity when various kinds of goal-oriented perceptual tasks are given). A plausible interpretation of default activity is that it represents non-task-related ‘thinking’ that involves structured combination of semantic representations (Binder et al., 1999). Many of the same areas associated with the default network are also associated with episodic memory retrieval, which also presumably requires structured combination of the stored individual semantic representations that are represented in the memory episode (Svoboda et al., 2006; Addis et al., 2007). Binder and colleagues have argued that this is why many studies do not show semantic effects in AG, because they are contrasted against this baseline of semantic combination. They also suggest that some semantic anomaly studies in fMRI incorrectly reported AG effects as posterior STG effects.

Binder and colleagues (e.g. Humphries et al., 2007) have suggested that AG may be involved in integrating new information into a context. However, so many kinds of semantic combination and so many different ways of implementing it are possible that I think it is difficult to make a strong argument about the particular function that AG is involved in at this point. Importantly though, future imaging studies looking at semantic processing should include a sufficiently demanding perceptual task as a baseline in order to alleviate concerns that semantic activity is being subtracted out semantic activity with the baseline in the form of non-task related thought.

4.6 Linking the anatomy to the electrophysiology

So far in this chapter I have set out functional hypotheses for several cortical areas which have been associated with semantic processing within sentences, based on previous neuroimaging and neuropsychological studies. In this section I turn to the question of how differential activity in these areas is reflected in the ERP or the MEG signal. In particular, several of these areas (anterior and middle inferior frontal cortex, anterior temporal cortex) have demonstrated differential activity for manipulations of contextual support in sentence processing, as I reviewed in Chapter 3. In Chapter 3, I provided evidence that differential activity in posterior middle temporal cortex is reflected in the N400 effect in ERP. Given our functional hypotheses about these other areas, is it likely that they also contribute to the N400 effect, or are they reflected in other parts of the ERP or not at all?

Left inferior frontal gyrus is the area that is most frequently associated with sentence context manipulations, with all three IFG areas showing significant effects in at least some studies, and this area also showed effects in long SOA semantic priming studies. Several studies using techniques such as MEG distributed source localization and event-related optical signal recordings during typical N400 semantic anomaly paradigms (Halgren et al., 2002; Tse et al., 2008) have localized differential activity in left inferior frontal cortex following earlier effects in left temporal cortex. These results suggest that IFG effects may be reflected later in the ERP waveform than the MTG effects. These effects may be reflected as either the later part of the N400 effect or as the 'post-N400 positivity' (Van Petten & Luka, 2006; Matsumoto et al., 2005; Federmeier et al., 2007).

The post-N400 positivity

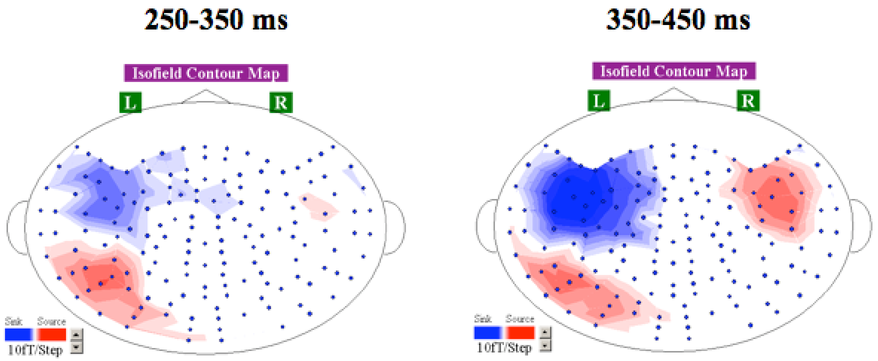
In many N400 studies, the incongruous condition produces an ERP that not only is more negative at the N400 peak, but also is more positive at the broad positive peak that follows. It is so far unclear why some studies find a late positivity and some do not (see Van Petten & Luka, 2006, for review) and whether this late positivity is related to the positivity observed in recent studies for certain kinds of thematic reversal anomalies (see Kuperberg, 2007, and Stroud, 2008, for discussion). Furthermore, some of the late positivities reported have a posterior focus and others appear stronger in frontal electrodes, which if real, suggests late effects with distinct functional correlates

Although relatively little is known about these late positivities, there are a few pieces of evidence suggesting that they may be associated with cases in which multiple candidates are competing with each other, as might be expected if this effect reflected differential activity in mid-IFG. First, a recent study that observed such a post-N400 positivity strongest in frontal electrodes showed that its amplitude is only increased for unexpected endings when they follow high-constraint (highly predictable) sentence contexts; no such effect is observed for corresponding low-constraint contexts that do not predict a specific word (Federmeier et al., 2007). Similarly, a semantic priming study showed a late frontal positivity for targets in unrelated relative to neutral prime contexts – the same contrast that showed a left posterior IFG effect in Gold et al.'s 2006 study (Holcomb, 1988). The same study showed that a late positivity was observed for unrelated relative to related targets under strategic priming conditions but not under automatic priming conditions, consistent with our observation that IFG effects are only observed for long SOA priming experiments. Van Petten and Luka (2006) note an ERP study of patients with frontal lesions who failed to show a post-N400 positivity (Swick, Kutas, & Knight, 1998). Finally, I have noticed anecdotally that late positivities appear to be especially pronounced in cases in which the anomalous word is semantically or phonologically related to the expected ending (Van Petten et al., 1999; van den Brink et al., 2001; Diaz & Swaab, 2006; Connolly & Phillips, 1994; Vissers, Chwilla, & Kolk, 2006; Federmeier & Kutas, 1999a), which might require more effortful selection.

While these results are consistent with the hypothesis that the late post-N400 positivity reflects differential IFG activity, the interpretation of these results is very much complicated by the fact that I have proposed that anterior and posterior LIFG carry out very different functions, both of which seem to be affected by contextual manipulations, according to the fMRI evidence. Research using MEG and EROS source localization is still too untested to reliably distinguish sources from such closely neighboring areas. What would be needed is to examine electrophysiological recordings for conditions in which fMRI dissociates anterior and posterior LIFG (or, following our functional hypothesis, conditions that dissociate retrieval demands from selectional demands). One such case is the contrast in directionality that Gold et al. (2006) show for anterior and posterior LIFG relative to a neutral-prime condition: anterior LIFG shows a relative reduction in activity for semantically related pairs, while posterior LIFG shows a relative increase in activity for unrelated pairs. An ERP study of semantic priming that included such a neutral condition showed an increased positivity for the unrelated condition, consistent with a posterior LIFG source (Holcomb, 1988). On the other hand, there does not appear to be a clear association between SOA for semantic priming and the late positivity; while one ERP study showed a hint of increased positivity in the long SOA case (Anderson & Holcomb, 1995), others did not (Deacon et al., 1999; Hill et al., 2002; Rossell et al., 2003). If posterior LIFG activity reflects selection between the predicted target and the actual input, on this account it should be reflected in the ERP to unrelated targets at long SOAs, which do not match the prediction due to the prime.

Second stage of N400 effect

Another possibility is that differential LIFG activity contributes to a multi-generator N400 effect. In Chapter 4, I provided evidence that left posterior MTG is a source of the N400 effect. However, the N400 effect is long-lasting, and may reflect the contributions of multiple cortical areas over time. Results from the MEG study I presented in Chapter 3 are consistent with previous MEG studies of the N400 effect (Halgren et al., 2002; Maess et al., 2006; Pylkkänen & McElree, 2007) in suggesting the existence of multiple generators across the interval in which the N400 effect is usually reported. In this study and in a subsequent replication (Lau, Almeida, AbdulSabur, Braun, & Poeppel, 2008) we found that in an early time-window (~250-350 ms), the N400 effect was reflected only as a left-hemisphere dipole, while in a later time-window (~350-500 ms), a cluster of right anterior sensors also showed a significant difference. For convenience, I repeat the field pattern from Experiment 1 as Figure 23 below.


250-350 ms

350-450 ms

Figure 23. Statistically thresholded grand-average whole-head topography from Experiment 1 for the sentence ending contrast (contextually unsupported – contextually supported) averaged across two time-windows chosen by visual inspection, showing only those sensors for which the difference between conditions was significant across participants (*p* < .01). In the first time-window, the two left-hemisphere clusters were larger than would have been expected by chance, while in the second time-window, an additional right-hemisphere cluster was also larger than would have been expected by chance.

The early pattern is well-explained by the left MTG generator I posited in Chapter 3, but what about the later pattern? One possibility is that this field pattern reflects the combination of a left middle temporal source and two lateral anterior cortical sources. This account would assume that, due to the orientation of the two anterior sources, and because MEG sensors are not positioned around the entire cortical 'sphere', only one pole of each dipole is captured by the sensors, much like early visual MEG responses such as the M170. On this scenario, bilateral inferior frontal or bilateral anterior temporal cortex would be the best candidates for generating the second half of the N400 effect.

Most of the language comprehension studies reviewed report effects in left anterior IFG only. However, a recent semantic anomaly fMRI study by Menenti and colleagues (2009) reported significant effects of anomaly of equivalent size in anterior IFG bilaterally. Another interesting point in support of this hypothesis is that differences in anterior IFG track differences in MTG across a number of studies (e.g. Buckner et al., 2000; Badre et al., 2005; Dobbins & Wagner, 2005; Gold et al., 2006; Staresina et al., 2009); this would be consistent with the N400 effect looking fairly similar across most paradigms even though on this account it is composed by more than one source. What would need to be explained is why so many studies using the same N400 paradigms as in ERP studies do not report bilateral effects in anterior IFG. The same question arises if the bilateral activity is attributed to anterior temporal cortex instead; as I review above, right ATC activity tends to be associated specifically with auditory comprehension.

Alternatively, Pylkkänen and colleagues (Pylkkänen & McElree, 2007; Brennan & Pylkkänen, 2008; Pylkkänen, Martin, McElree, & Smart, 2008) have argued that a bilateral response with similar timing topographic field pattern during the N400 time interval does not reflect two lateral sources, but rather one ventromedial PFC source. Since vmPFC has frequently been implicated in studies of episodic and autobiographical memory and imagining the future (e.g. Svoboda, McKinnon, & Levine, 2006; Hassabis & Maguire, 2009) and activation of contextual frames (Bar, 2009; Aminoff, Schacter, & Bar, 2008), increased activity in sentences containing pragmatic violations could perhaps reflect access of stored contextual frames to make sense of the sentence. However, this area also was not frequently reported in the N400 studies in fMRI that I reviewed in the meta-analysis.

At the moment, then, there is suggestive evidence that the N400 effect may reflect other generators in addition to MTG, particularly in the latter part of the N400 time interval, but this evidence does not converge in implicating one region. Future MEG studies may help constrain hypotheses about this second part of the N400 effect by testing whether manipulations focusing on particular processes such as selection or integration succeed in selectively affecting only this later part of the effect.

4.7 Conclusion

In Chapter 3 I presented evidence that posterior middle temporal cortex is involved in storage of lexical-semantic information. In this chapter I have examined the contribution of other areas implicated by contextual manipulations in language processing. I summarized evidence that left aIFG is involved in non-automatic retrieval of semantic information (Badre & Wagner, 2007) and I proposed that this function makes it possible for this region to instantiate predictive pre-activation of stored representations in posterior temporal cortex. I also reviewed evidence that left mIFG is involved in selection among activated representations, that left ATC may be involved in the combination or representation of structured relationships between words in the sentence, and that left AG may be involved in the combination or representation of conceptual representations. Finally, I discussed how the neurophysiological effects observed in ERP and MEG measures might map onto these regions of interest.

In Figure 24, I provide a rough suggestion of how information might flow between these regions during the course of processing a word within a sentence context. aIFG is involved in pre-activating representations stored in MTG based on the prior context. When the auditory or visual information associated with the critical word is presented and processed by modality-specific areas, this bottom-up information also impacts the state of representations at MTG. mIFG is involved in the process of selecting one of the activated lexical candidates at MTG to serve as input for computations involved in updating and maintaining representations of the sentence and discourse that may be supported by ATC and AG. Information from these updated representations of the context are used to pre-activate candidate representations at MTG for the next step.

This picture simply represents the basic information flow necessary to support the operations with which BOLD activity in these regions is associated with experimentally, and thus should not be taken as a strong claim about what areas of cortex actually do the computations and by what neural pathways they pass information to each other. Although this framework is very preliminary, it generates specific anatomical hypotheses that can be tested more systematically in the future.

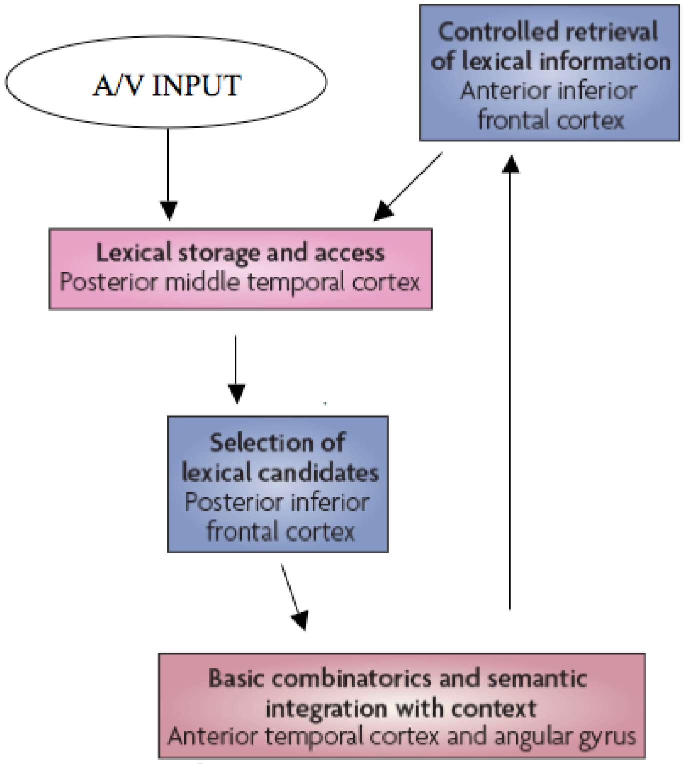


Figure 24. Schematic illustration of information flow that would be required if the cortical regions discussed in Chapter 5 fulfill the functions proposed for them in the processing of words in context.

5 Syntactic predictions I: Mechanism and timecourse

5.1 Introduction

In the previous three chapters I have focused on mechanisms of top-down facilitation of lexical access by sentence context, presenting new evidence for such mechanisms and developing a model of how these mechanisms may be instantiated neuroanatomically. In the next two chapters I will turn to top-down mechanisms for syntactic processing, which raise different issues. While there is a huge literature on effects of predictive context in lexical access and a debate about whether predictive top-down facilitation actually occurs, there is a much smaller literature on effects of predictive context in syntactic processing.

In this chapter I will discuss a study that makes use of an ERP component traditionally associated with syntactic violations—the early left anterior negativity (ELAN)—to try to provide evidence for consequences of syntactic prediction on online processing. Although the results of this experiment do not provide definitive evidence for syntactic prediction, they provide a good starting point for beginning to think about how syntactic predictions might be implemented, how the timing of contextual effects bears on predictive interpretations, and how future studies can be best designed to further investigate syntactic prediction.

5.2 Syntactic prediction

In this section I will examine what evidence exists for syntactic prediction during processing. By syntactic prediction I mean predictions that depend on only the syntactic information encoded in the preceding context.

One way in which syntactic predictions might facilitate processing of bottom-up input is through pre-activation of lexical items of a syntactic category that appears in all possible expansions of the current node or which are likely to appear in the expansion of the current node based on previous experience. For example, when a determiner like *the* is encountered, all of the nouns in the lexicon might be pre-activated. Pre-activation could also extend to lower representational levels, as recent work by Farmer and colleagues argues that distinct statistical regularities of phonological composition can be observed for different syntactic categories are associated (Farmer, Christiansen, & Monaghan, 2006); thus, a syntactic prediction could lead to pre-activation of particular phonological or orthographic representations (Dikker et al., submitted). However, because most syntactic predictions are not locked to the next position in time (e.g., even though a determiner predicts a noun, the next word could be an adjective or even a degree adverb, *the very blue sheep*) and because many syntactic categories contain so many members, it is unclear whether syntactic pre-activation would measurably facilitate lexical access except in a few constrained cases (Tanenhaus and Lucas, 1987).

Some evidence supporting such a mechanism comes from studies showing that lexical decisions are faster for words of a syntactic category predicted by the context than for words of a syntactic category not predicted by the context (e.g., Lukatela et al., 1982; Lukatela et al., 1983; Wright & Garrett, 1984; Randall & Marslen-Wilson, 1998; Bölte and Connine, 2004). In the examples below (9) from Wright and Garrett (1984), none of the targets are semantically congruent, but they vary in syntactic congruency. Reaction times were consistently longer for the targets that did not match the syntactic prediction.

(9)
a. If your bicycle is taken, you must FORMULATE (*verb context syn. congr.)*
b. If your bicycle is taken, you must BATTERIES (*verb context syn. incongr.)*
c. For now the happy family lives with FORMULATE (*noun context syn. incongr.)*
d. For now the happy family lives with BATTERIES (*noun context syn. congr.)*

If it is assumed that lexical decision speed primarily reflects speed of lexical access, then these results support the hypothesis that syntactic prediction allows facilitative pre-activation of words of a particular syntactic category. However, one problem with this interpretation is that participants probably automatically attempted to integrate the target with the rest of the sentence context, even though it is not required by the lexical decision task. Because the unpredicted conditions are also the syntactically incongruent ones, the delayed reaction times may simply reflect a response to incongruity rather than a predictive mechanism.

Another way in which syntactic knowledge might facilitate processing of the bottom-up input is that the grammar may limit the number of possible attachment sites for the current input more in certain syntactic contexts than others. Depending on the parsing algorithm, this could speed attachment decisions in grammatical sentences (Staub & Clifton, 2006; Yoshida & Sturt, 2009), but furthermore, in ungrammatical sentences, strongly constraining syntactic contexts may allow the parser to realize faster that there is no grammatical attachment site according to the current structural analysis being pursued.

In the following study, we considered whether the second possibility might explain how a response to a certain class of syntactic violations has been observed so early in processing with ERP. If this early ERP effect can be shown to reflect the effects of syntactic prediction, determining the properties of this effect can provide us with a better model of how syntactic constraint facilitates processing.

5.3 The ELAN component

In the early 1990s, Neville and colleagues (Neville et al., 1991) and Friederici and colleagues (Friederici, Pfeifer, & Hahne, 1993) showed that grammatical category violations like those shown in (10) and (11) elicited an increased left anterior negativity with a latency of 100-250 ms in the ERP, relative to the comparable grammatical sequences shown (Figure 25). This early response component came to be known as the Early Left Anterior Negativity (ELAN).

(10)
a. *The scientist criticized Max's of proof the theorem.
b. The scientist criticized Max's proof of the theorem.

(11)
a. *Die Kuh wurde im gefüttert.
The cow was in-the fed.
b. Die Kuh wurde im Stall gefüttert.
The cow was in-the barn fed.

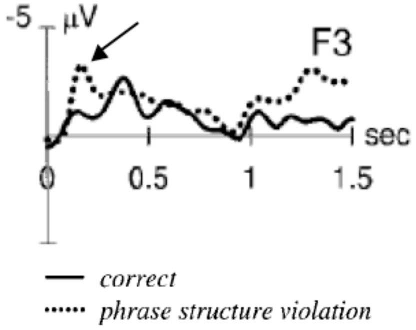


Figure 25. Illustration of an ELAN response to auditorily presented phrase structure violation from Hahne and Friederici (1999). This study contrasted the response to sentences of the sort presented in (11).

The ELAN response seems to be elicited only in very specific contexts, usually the syntactic sequences in (10) and (11), but also in variants upon them in Spanish (Hinojosa, Martin-Loeches, Casado, Muñoz, & Rubia, 2003), French (Isel, Hahne, Maess & Friederici, 2007), and German (Rossi, Gugler, Hahne, & Friederici, 2005). Many types of ungrammaticality do not elicit an ELAN response, such as inflection/agreement violations, case violations, and violations of wh-fronting, which sometimes show a later left anterior negativity (LAN) in the 300-500 ms range (e.g., Coulson, King, & Kutas, 1998; Friederici et al., 1993; Gunter, Stowe, & Mulder, 1997; Kaan, 2002; Kutas & Hillyard, 1983; Münte, Matzke, & Johannes, 1997; Osterhout & Mobley, 1995; Kluender & Kutas, 1993a, 1993b), and syntactic garden paths (temporary ungrammaticalities) and subcategorization violations, which elicit a late positivity known as the P600 and sometimes a LAN (Kaan & Swaab, 2003; Osterhout, Holcomb, & Swinney, 1994; Friederici, Hahne, & Mecklinger, 1996; Ainsworth-Darnell, Shulman, & Boland, 1998). The ELAN also appears not to be sensitive to task manipulations (Hahne & Friederici, 2002) or to the experiment-wide probability of the violation (Hahne & Friederici, 1999).

The earliness of the ELAN response was what initially suggested to us that the response might reflect violation of syntactic prediction rather than simply recognition of the ungrammaticality of the continuation. There is some variation in the latency of the response, but the onset is in the 100-200 ms range for a number of studies, including the original report in English reading by Neville et al. (1991) and most of the German auditory-based studies by Friederici and colleagues in which the offending participle is clearly marked by the participial prefix *ge-* and the suffix *-t* (e.g., Friederici et al., 1993; Hahne & Friederici, 1999). Given the time required for basic perceptual processing on the input—at least 60 ms for visual word input on a conservative estimate (Sereno & Rayner, 2003)—this time-range falls near the earliest time-range that has been proposed for lexical access processes (Sereno, Rayner, & Posner, 1998; Allopenna, Magnuson, & Tanenhaus, 1998; van Petten, Coulson, Rubin, Plante, & Parks, 1999; Hauk et al., 2006), suggesting that it is perhaps too early to reflect integration difficulty. On the other hand, if prediction allows processing to begin before the word, it could explain the earliness of the ELAN response.

However, as I will return to in the discussion at the end of the chapter, making an argument for prediction from timing alone is tricky given the lack of consensus about the timing of access and integration mechanisms. Friederici has argued that the ELAN, despite its earliness, does reflect difficulty of integrating the incoming word with the current structure rather than violation of prediction (see Friederici, 2002, for review). On her hypothesis, the ELAN is a response elicited by any incoming word whose grammatical category cannot be integrated into any possible elaboration of the existing phrase structure. For Friederici, the response is so rapid not because it reflects an earlier access stage, but because access to syntactic category information is faster than any other kind of lexical information, and therefore integration of this information with the previous material can happen very quickly.

We conducted an ERP experiment designed to test more directly our hypothesis that the ELAN response is so early because it reflects syntactic prediction. Previous ERP experiments have mainly contrasted ungrammatical sentences with their grammatical counterparts. In this experiment, we focused on the contrast between two well-matched ungrammatical sentences that differed in the strength of the syntactic category prediction induced by the context at the critical word, allowing us to pull apart integration difficulty from violation of predictions. This experiment was conducted in collaboration with Clare Stroud, Silke Plesch, and Colin Phillips (Lau, Stroud, Plesch, & Phillips, 2006).

5.4 Experiment 3

In this study, we used the possibility of ellipsis to manipulate whether a possessor at the left edge of an NP created a strong prediction for an overt noun. We did this by adding a preceding clause to the constructions that were previously used to elicit the ELAN in English. This allowed us to weaken the prediction for an overt noun after the possessor, as shown in 0.

(12) I don't like Bill's friend but I do like Max's ___.

Although normally a sentence without a noun following a possessor like *Max's* would be ungrammatical, in a sentence like 0 that allows ellipsis of the noun, the sentence can end grammatically with *Max's*, as the parallel structure in the first clause licenses the absence of the phonological material in the object NP of the second clause (Lobeck, 1995). Therefore, the ellipsis context may weaken the expectation at the possessor that an overt noun will follow. However, even though an overt noun is no longer required, a preposition following the possessor is still ungrammatical, as shown in 0.

(13) *I don't like Bill's friend but I do like Max's of six years.

By comparing the response to ungrammatical possessor-preposition sequences in ellipsis and non-ellipsis contexts, we can test whether the ELAN response is sensitive only to phrase structure violations per se, or whether it is sensitive to the strength of the syntactic category prediction at a particular position. If, as Friederici (2002) has suggested, the ELAN reflects simply the recognition that the syntactic category of the input word is incompatible with the previous material, both ungrammatical conditions should show an ELAN of equal size. However, if the ELAN reflects some process associated with violating a syntactic prediction, it may be larger in the non-ellipsis case where the prediction for an NP following the possessor is stronger, relative to the ellipsis case where the prediction is weaker.

Participants

Forty-one members of the University of Maryland community participated in the ERP study. None had participated in testing of materials (described below). Data from six participants were excluded due to technical problems and data from three participants were excluded due to high levels of artifacts in the EEG recordings. All 32 remaining participants (20 female; mean age 20.7; range 18–37 years) were healthy, monolingual native speakers of English with normal or corrected to normal vision, and all were classified as strongly right-handed based on the Edinburgh handedness inventory (Oldfield, 1971). All participants gave informed consent and were paid $10/hour for their participation, which lasted around 2.5 hours, including set-up time.

Materials

To confirm that our contextual manipulation successfully modulated the strength of the prediction for an overt noun following a possessor we conducted a preliminary offline sentence completion study. Fourteen participants from the University of Maryland community were asked to supply completions to fragments like (14) and (15), corresponding to the first 10 words of the syntactically incorrect experimental conditions, examples of which are presented in Table 7.

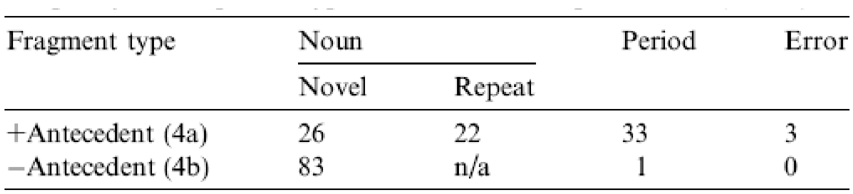
+ Ellipsis	Grammatical	Although Erica kissed Mary's mother, she did not kiss the daughter of the bride.
	Ungrammatical	*Although Erica kissed Mary's mother, she did not kiss Dana's of the bride.
- Ellipsis	Grammatical	Although the bridesmaid kissed Mary, she did not kiss the daughter of the bride.
	Ungrammatical	*Although the bridesmaid kissed Mary, she did not kiss Dana's of the bride.
Agreement	+ Agree	Although Matt followed the directions closely, he had trouble finding the theater.
	- Agree	*Although Matt follow the directions closely, he had trouble finding the theater.

Table 7. Sample set of conditions used for Experiment 3.

The +antecedent condition (14) contained a noun phrase with a possessor in direct object position of the first clause, which served as a potential antecedent for ellipsis in the second clause and allowed the string to stand complete without further elaboration. The -antecedent condition (15) served as the baseline condition, since the first clause contained no antecedent for ellipsis, and therefore the fragment had to be completed with a noun. Participants were instructed to insert suitable words to complete the sentence naturally, and were told that they could insert a period if they felt that no additional words were needed to complete the sentence. Six items were presented from each of the two conditions, within a questionnaire that included 48 filler items, all of which used a similar two-clause form.

(14) Although Erica kissed Mary’s mother, she did not kiss Dana’s. . .
(15) Although the woman kissed Mary, she did not kiss Dana’s. . .

A summary of results from the completion study is presented in Table 8. The results showed that there was a significant difference in the pattern of completions for the two conditions ($\chi 2 = 37.759$, $p < .001$). As expected, in the -antecedent condition, participants supplied a grammatical continuation with an overt noun in 99% of trials. In contrast, in 39% of the +antecedent trials participants simply ended the sentence with a period, indicating an elliptical interpretation for the object of the second clause. Additionally, in 26% of +antecedent trials participants supplied a noun that matched the noun in the first clause, suggesting that they constructed the same interpretation used in the elliptical completions. These results provide evidence that the contextual manipulation used in the ERP study is indeed able to weaken the prediction for a novel noun following the possessor, although it remains likely that a novel noun was still predicted on some proportion of the +ellipsis trials in the ERP study.



Fragment type	Noun		Period	Error
	Novel	Repeat		
+Antecedent (4a)	26	22	33	3
−Antecedent (4b)	83	n/a	1	0

Table 8. Frequency of completion types in the offline completion task (*n* = 14).

The materials for the ERP study consisted of sentence quadruples organized in a 2 × 2 factorial design and corresponding to the first four conditions in Table 7, plus pairs of sentences corresponding to the grammatical and ungrammatical agreement conditions. In all four category-violation conditions, the sentences consisted of a subordinate clause followed by a main clause; within each set, the two factors varied were the availability of ellipsis (presence vs. absence of a possessive NP in the first clause) and grammaticality (full NP vs. possessor preceding the preposition in the second clause). To facilitate a contrastive interpretation of the two clauses and thereby favor ellipsis, the second clause always contained the same verb as the first clause, and the second clause showed the opposite polarity of the first clause, i.e., the second clause was negated. To balance the number of words and the number of common nouns in the first clause across conditions, the subject of the first clause was a proper name in the +ellipsis conditions and a two word NP in the -ellipsis conditions. Within each level of the grammaticality factor, the sequence of words preceding the critical word preposition was identical across a span of at least five words.

The common noun in the second clause (e.g., *daughter*) was chosen to freely allow a following argument or modifier prepositional phrase (PP) headed by *of*. On the other hand, the choice of common noun in the possessive NP in the first clause was constrained in order to resist combination with the preposition *of*, and thereby to ensure that the possessor + *of* sequence in the second clause would be equally unacceptable in the two grammatically incorrect conditions. Some speakers of English find it marginally acceptable to elide the head noun to create a possessor + *of* sequence when an appropriate antecedent is available and the possessor may be construed as the agent of the event denoted by the noun (16). In cases where the possessor has a true genitive interpretation this ellipsis is fully impossible (17).

(16)
  a. ?John's description of the crime and Mary's of the leading suspects.
  b. ?John's news of the earthquake and Mary's of the relief effort.
  c. ??The barbarians' destruction of the city and the peasants' of the countryside.

(17)
  a. *John's vase of crystal and Mary's of solid silver.
  b. *Manchester United's director of coaching and Chelsea's of marketing.
  c. *Sue's friend of 15 years, and Sally's of six months.

One hundred and twenty eight sets of items for the category violation conditions were distributed across four presentation lists in a Latin Square design, such that each list contained 32 items per condition. In addition, 64 pairs of agreement items were distributed across two presentation lists in a Latin Square design. Each list of category violation items was combined with one of the agreement lists and 192 filler items to create four lists with 384 items each. The filler items were similar to the experimental items in maintaining a subordinate–main clause format, and were all grammatically correct. Thus, items from the four category violation conditions were outnumbered 2:1 by other items, and the ratio of grammatical to ungrammatical sentences in the experiment was 3:1. Furthermore, since the grammatical violations occurred in either the first clause (agreement conditions) or the second clause (category prediction conditions), participants needed to pay attention to the entire sentence in order to accurately judge the well-formedness of the sentence.

Procedure

Participants were comfortably seated in a dimly lit testing room around 100 cm from a computer monitor. Sentences were presented one word at a time in black letters on a white screen in 30 pt font. Each sentence was preceded by a fixation cross. Participants pressed a button to initiate presentation of the sentence, which began 1000 ms later. Each word appeared on the screen for 300 ms, followed by 200 ms of blank screen. The last word of each sentence was marked with a period, and 1000 ms later a question mark prompt appeared on the screen. Participants were instructed to read the sentences carefully without blinking and to indicate with a button press whether the sentence was an acceptable sentence of English. Feedback was provided for incorrect responses. This task is similar to tasks used in previous studies of responses to category violations. Each experimental session was preceded by a 12 trial practice session that included both grammatical and ungrammatical sentences. Participants received feedback and were able to ask clarification questions about the task at this time. The experimental session itself was divided into six blocks of 64 sentences each.

EEG recording

EEG was recorded from 30 Ag/AgCl electrodes, mounted in an electrode cap (Electrocap International): midline - Fz, FCz, Cz, CPz, Pz, Oz; lateral - FP1/2, F3/4, F7/8, FC3/4, FT7/8, C3/4, T7/8, CP3/4, TP7/8, P4/5, P7/8, and O1/2. Recordings were referenced online to the linked average of the left and right mastoids, and re-referenced offline to the common average reference, as discussed further below. Additional electrodes were placed on the left and right outer canthus, and above and below the left eye to monitor eye movements. EEG and EOG recordings were amplified and sampled at 1 kHz using an analog bandpass filter of 0.1–70 Hz. Impedances were kept below 5 kΩ.

EEG analysis

All comparisons were made based upon single word epochs, consisting of the 100 ms preceding and the 1000 ms following the critical words. Epochs with ocular and other large artifacts were rejected from analysis based on visual screening. This affected 10.9% of trials, ranging between 9.1% and 13.8% across conditions. The waveforms of the individual trials were normalized using a 100 ms pre-stimulus baseline. Averaged waveforms were filtered offline using a 10 Hz low-pass filter for presentation purposes, but all statistics are based on unfiltered data. The following latency intervals were chosen for analysis, based on the intervals used in the literature and on visual inspection: 0–200 ms, 200–400 ms (ELAN), 300–500 ms (LAN), and 600–1000 ms (P600).

To test for lateral effects, four quadrants of electrodes were defined as follows: left anterior (F7, FT7, F3, and FC3), right anterior (F4, FC4, F8, and FT8), left posterior (TP7, P7, CP3, and P3), and right posterior (CP4, P4, TP8, and P8). ANOVAs were performed hierarchically, using the within-subjects factors ellipsis (ellipsis available/unavailable), grammaticality (grammatical/ungrammatical), hemisphere (left/right), anteriority (anterior/posterior), and electrode (four per region). In addition, ANOVAs were performed separately on the midline electrodes, with two regions, anterior (Fz, FCz, and Cz) and posterior (CPz, Pz, and Oz), and the factors ellipsis, grammaticality and anteriority. All p-values reported below reflect the application of the Greenhouse-Geisser correction where appropriate, to control for violations of the sphericity assumption (Greenhouse & Geisser, 1959), together with the original degrees of freedom. Due to the large number of possible interactions in this design, we report as significant only those interactions for which follow-up analyses yielded significant contrasts within the levels of the interacting factors.

Behavioral results

Overall accuracy on the behavioral grammaticality judgment task for the four category violation conditions was 95.4%. The agreement conditions showed an overall accuracy that was somewhat lower, 88.4%. Average accuracy for the ungrammatical agreement condition was only 82.5%, suggesting that participants either experienced some difficulty in detecting these violations, or were less attentive to these violations, perhaps because they were less striking than the category violations or because they appeared earlier in the sentence.

ERP Results: pre-critical word

Visual inspection of the responses to the grammatical and ungrammatical conditions suggested that the waveforms already diverged at the word preceding the critical word *of*. This is perhaps not surprising in light of the lexical differences between conditions in this position. Pre-existing differences in the responses to grammatical and ungrammatical conditions could bias the analysis of responses to the preposition *of*. To statistically test for pre-existing differences, an ANOVA was performed on the 300–500 ms interval following presentation of the word preceding *of*, which falls immediately before the presentation of the critical word. At this interval the ungrammatical conditions were more negative at anterior scalp regions and slightly more positive at posterior regions, relative to the grammatical conditions. This yielded a main effect of grammaticality ($F(1, 31) = 10.62$, $p < .01$), and interactions between grammaticality and hemisphere ($F(1, 31) = 4.25$, $p < .05$) and between grammaticality and anteriority ($F(1, 31) = 18.41$, $p < .001$). The ANOVAs comparing grammatical and ungrammatical conditions within each level of the ellipsis factor showed similar patterns within both the +ellipsis and -ellipsis pairs. The ungrammatical +ellipsis condition showed a more negative response than the grammatical +ellipsis condition at anterior scalp regions, including the left anterior quadrant ($F(1, 31) = 11.52$, $p < .01$), the right anterior quadrant ($F(1, 31) = 4.16$, $p < .06$), and the anterior midline region, ($F(1, 31) = 14.88$, $p < .01$), but showed a more positive response than the grammatical condition in the right posterior quadrant ($F(1, 31) = 7.52$, $p < .05$). The response to the -ellipsis ungrammatical condition was more negative than the response to the -ellipsis grammatical condition in the left anterior quadrant, ($F(1, 31) = 14.61$, $p < .01$), and in the anterior midline region, ($F(1, 31)$

= 13.74, $p < .01$). The effects of grammaticality are unsurprising, since the comparison involves different words with different grammatical roles: in the grammatical condition the word before *of* is a noun (e.g., *director*), while in the ungrammatical condition the word before *of* is a possessor (e.g., *John's*) that precedes the anticipated noun. Due to the differences in ERP responses at the preceding word, it was difficult to establish a reliable baseline interval or to reliably identify effects of the processing of the word *of*. Therefore, in what follows we only report comparisons within each level of the grammaticality factor. Although this means we cannot directly evaluate the effect of grammaticality, the advantage of this approach is that all comparisons are based on conditions that are lexically very well matched and that are also identical through at least the five regions preceding the critical word *of*.

In addition to the effect of grammaticality, the ANOVA revealed that in the 300–500 ms interval after the word preceding the critical word of there was a main effect of ellipsis ($F(1, 31) = 7.69$, $p < .01$), with no significant or marginally significant interactions between ellipsis and any other factors. However, the difference in amplitude was small, with the average response to +ellipsis conditions being 0.12 μV more negative than the response to -ellipsis conditions. Visual inspection suggested that this effect arose from a difference between the grammatically correct +ellipsis and -ellipsis conditions, and that the two ungrammatical sentences were closely matched. The statistics confirmed this: an ANOVA comparing the grammatical +ellipsis and -ellipsis grammatical conditions yielded a significant effect of ellipsis ($F(1, 31) = 9.13$, $p < .01$), with no significant or marginally significant interactions between ellipsis and any other factor. Again, the amplitude difference was small; the response to the grammatical +ellipsis condition was 0.17 μV more negative than the response to the grammatical -ellipsis condition. This effect is a concern since effects of the ellipsis manipulation were not anticipated at this point in the sentence. However, since the amplitude difference was small and non-focal, relative to the approximately 1 μV focal ELAN effect at the critical word, we assume that the difference at the pre-critical word in the grammatical conditions does not compromise our main conclusions. Importantly for the main topic of our study, there was no difference between the ungrammatical +ellipsis and the ungrammatical -ellipsis conditions at the pre-critical word.

ERP Results: critical word

Due to pre-existing differences in the grammaticality factor at the pre-critical word, as discussed above, subsequent analyses focus on comparisons within each level of the grammaticality factor. In the comparison of responses to the word *of* in the two ungrammatical conditions (Figure 26 and Table 9), there was an effect of ellipsis in the 0– 200 ms interval; the +ellipsis condition showed a slightly greater negativity than the -ellipsis condition, although this difference was not significant in any individual region. In the 200–400 ms interval, the interval in which we expected to see modulation of the ELAN effect, the response to the -ellipsis condition in the left anterior scalp quadrant was significantly more negative than in the +ellipsis condition. There was also a significant effect of ellipsis in the right posterior quadrant, due to a slightly more negative response to the +ellipsis condition than to the -ellipsis condition. In the 600–1000 ms window there were no differences that would be characteristic of a P600. Although this could in principle reflect a lack of P600 in both conditions, it seems more likely that the match between conditions reflects an equivalent late positive effect in both conditions.

	0–200 ms	200–400 ms	300–500 ms	600–1000 ms
Overall ANOVA—four quadrants				
Ellip (1,31)	5.11*	—	—	—
Ellip x lat (1,31)	—	7.60*	5.43*	4.95*
Ellip x ant (1,31)	—	—	—	—
Ellip x lat x ant (1,31)	—	—	3.49**	—
Overall ANOVA—midline				
Ellip (1,31)	—	—	—	—
Ellip x ant (1,31)	—	—	—	—
Anterior regions				
Left Ellip (1,31)	—	6.60*	3.68**	—
Mid Ellip (1,31)	—	—	—	—
Right Ellip (1,31)	—	—	—	—
Posterior regions				
Left ellip (1,31)	—	—	—	—
Mid ellip (1,31)	—	—	—	—
Right ellip (1,31)	—	5.85*	2.90**	—

* <.05.
** <.1.

Table 9. ANOVA F-values at the critical word *of* for the comparison between the +ellipsis and –ellipsis ungrammatical conditions.

Note that the analyses presented in this study are based upon data that used a common average reference, in which individual electrodes are referenced to the average of all electrode voltages. Although electrophysiological studies of sentence comprehension have by convention used a mastoid reference, the average reference is widely used in some areas of ERP research. A potential drawback to using the mastoid reference is that this method assumes that the mastoids do not pick up any electrical activity that correlates with the experimental measure, because this activity is subtracted from the voltages measured at all electrodes (Dien, 1998). Use of an average reference here avoided this potential confound. We found that the comparison of the ungrammatical -ellipsis and +ellipsis conditions shows a more negative response at left anterior electrodes than at other regions for either choice of reference.

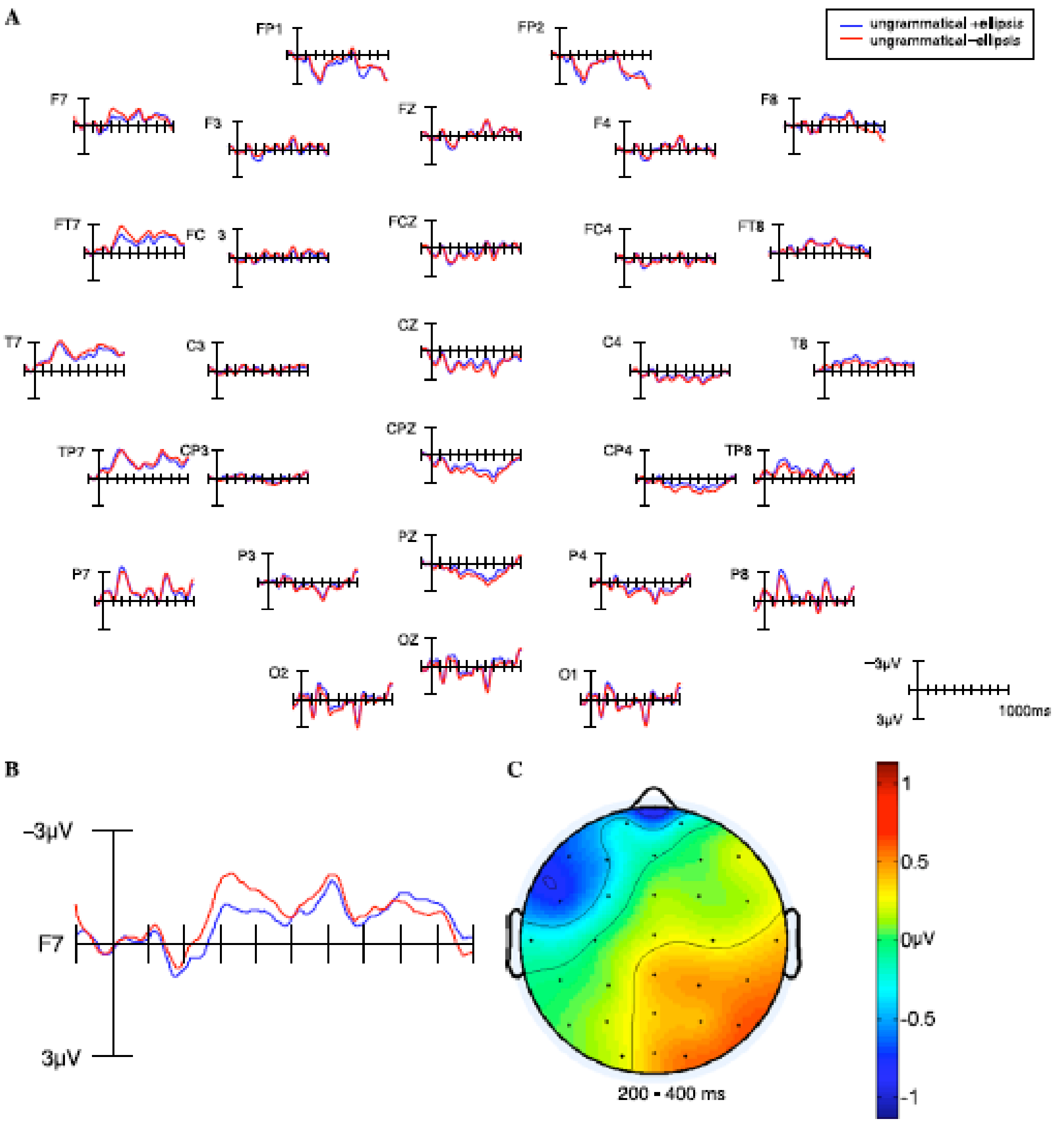
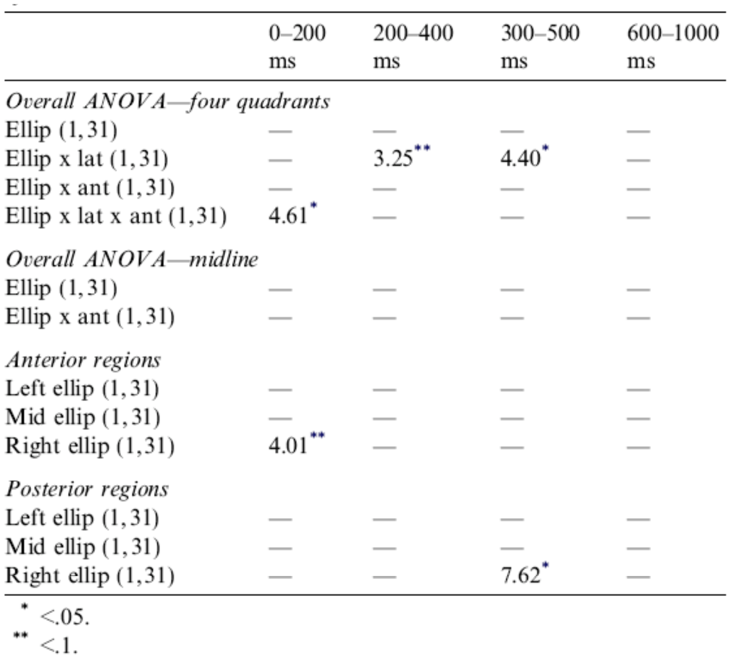


Figure 26. Grand-average ERPs in the +ellipsis ungrammatical (blue) and –ellipsis ungrammatical (red) conditions, computed using an average reference, showing the waveforms at (A) all electrodes, and (B) left anterior electrode F7. (C) presents a topographic plot of the average difference between the two conditions (-ellipsis ungrammatical - +ellipsis ungrammatical) across the scalp in the 200-400 ms time-window following the critical word.

We predicted that manipulation of ellipsis should not affect responses to the word *of* in the grammatical conditions, since the two conditions were identical in the preceding five words and since the contextual manipulation did not affect the acceptability of these conditions or the possible anticipated continuations at the word preceding *of*. The relevant results are shown in Figure 27 and Table 10.

Responses to the word *of* in the grammatically correct conditions showed a significant three-way interaction of ellipsis with hemisphere and anteriority in the 0–200 ms interval. This interaction was due to a marginally significant effect of ellipsis at the right anterior region, due to a small negativity (0.14 µV) in the -ellipsis grammatical condition. In the 200–400 ms interval there was a marginally significant interaction between ellipsis and hemisphere, but the effect of ellipsis was neither significant nor marginally significant at any individual region, suggesting that the effect was spurious. In the 300–500 ms interval, the +ellipsis grammatical condition showed a more negative response than the -ellipsis grammatical condition in the right posterior quadrant. There were no other significant effects of ellipsis at any quadrant in any time interval. The scarcity of reliable contrasts in the grammatical conditions suggests that the manipulation of ellipsis in the first clause did not appreciably affect the reading of the word of in the second clause in the two grammatical conditions.



	0–200 ms	200–400 ms	300–500 ms	600–1000 ms
Overall ANOVA—four quadrants				
Ellip (1,31)	—	—	—	—
Ellip x lat (1,31)	—	3.25**	4.40*	—
Ellip x ant (1,31)	—	—	—	—
Ellip x lat x ant (1,31)	4.61*	—	—	—
Overall ANOVA—midline				
Ellip (1,31)	—	—	—	—
Ellip x ant (1,31)	—	—	—	—
Anterior regions				
Left ellip (1,31)	—	—	—	—
Mid ellip (1,31)	—	—	—	—
Right ellip (1,31)	4.01**	—	—	—
Posterior regions				
Left ellip (1,31)	—	—	—	—
Mid ellip (1,31)	—	—	—	—
Right ellip (1,31)	—	—	7.62*	—

* <.05.

** <.1.

Table 10. ANOVA F-values at the critical word *of* for the comparison between the +ellipsis and –ellipsis grammatical conditions.

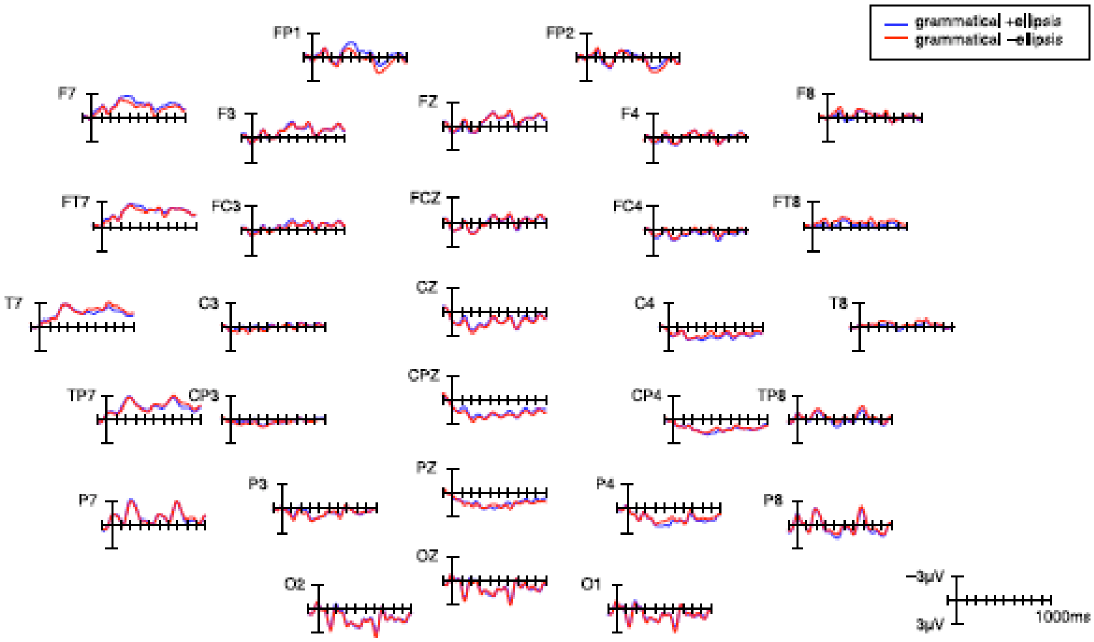


Figure 27. Grand-average ERPs in the +ellipsis grammatical (blue) and –ellipsis grammatical (red) conditions, computed using an average reference.

ERP Results: Agreement violation (control)

For the comparison between the conditions in which subject-verb agreement was manipulated, there was no effect of grammaticality in any of the early intervals: 0–200 and 200–400 ms (Figure 28). Nor was there any significant effect in the LAN interval of 300–500 ms. In the 600–1000 ms interval, there was a pattern of a posterior positivity and a corresponding anterior negativity, a pattern that is the average reference counterpart of the posterior positivity observed when using a mastoid reference. The response to the agreement violation had the broad central-posterior scalp distribution typical of the P600 response to syntactic anomaly.

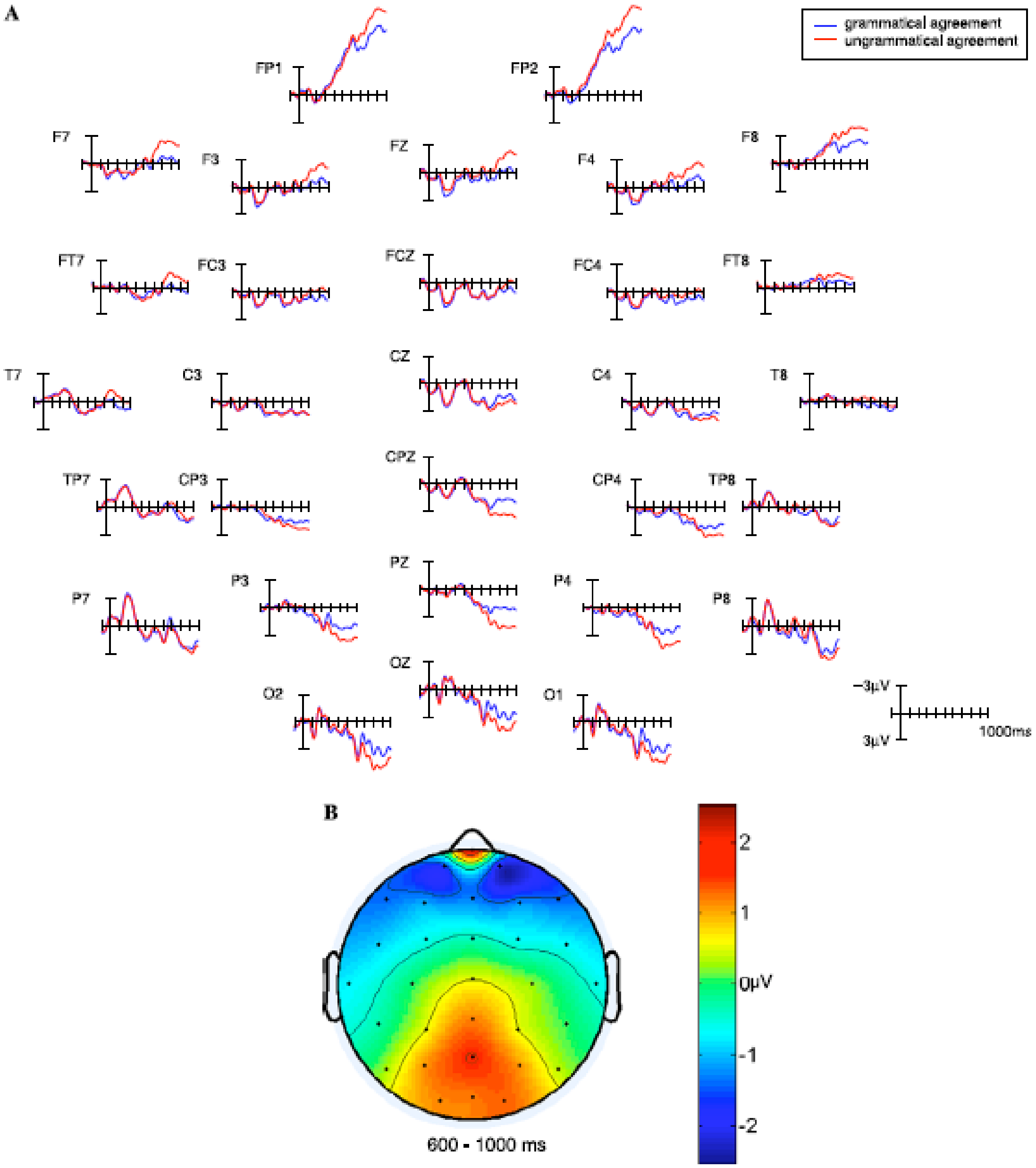


Figure 28. Grand-average ERPs in the grammatical (blue) and ungrammatical (red) subject-verb agreement conditions, computed using an average reference. (A) presents the waveforms for both conditions at all electrodes, and (B) presents a topographic plot of the average difference between conditions (ungrammatical agreement – grammatical agreement) in the 600-1000 ms time-window.

Discussion

The results of this experiment demonstrate that a phrase structure violation following a context which has unfulfilled syntactic and semantic requirements (a ‘predictive’ context) is processed differently than a phrase structure violation following a context that does not have unfulfilled requirements (an ‘unpredictive’ context). This difference took the form of an increased early left anterior negativity and right posterior positivity at around 200 ms in the ERP for ungrammatical continuations in strongly predictive contexts relative to those in weakly predictive contexts. In this comparison all words in the critical clause were identical and the violation was the same, but the response to the violation differed depending on the strength of constraint provided by the context.

We were able to bypass some of the problems of interpretation discussed in Section 5.2 by combining ERP, a technique with good temporal resolution, with a novel design that varied the strength of syntactic prediction while controlling both ungrammaticality and the local context of the ungrammatical word. In previous ERP designs, ungrammaticality and violation of prediction tracked each other, such that their effects could not be dissociated. Furthermore, most experiments that showed ELAN effects used materials that contained lexical differences at the region immediately preceding the critical word, and this made it difficult to exclude the possibility that the early differences on the critical word were actually due to differences in the response to the previous word. Conversely, reaction times reflect both early and late processes in one measure and may well have been dominated by the effect of ungrammaticality in a behavioral version of the ellipsis design. Therefore, it was likely the combination of the design and technique together that allowed us to show that the effect observed was more specific than a simple effect of ungrammaticality.

These findings clearly cannot be explained by Friederici's (2002) hypothesis that the ELAN indexes only the incompatibility of the syntactic category of the current word with the previous structure. I will next detail three possible interpretations of the ELAN that are consistent with the current findings: 1) the ELAN directly indexes violation of prediction; 2) the ELAN indexes early recognition of ungrammaticality, which is only made possible by syntactic prediction; 3) the ELAN indexes the response to a certain type of ungrammaticality, the failure to fulfill grammatical requirements of the previous input, and does not necessarily reflect prediction.

On the first account, the ELAN reflects a simple 'mismatch' response to the inconsistency between the syntactic category that is predicted and the syntactic category of the word that is presented. Any incoming input that does not match the predicted category automatically elicits an increased ELAN response. In this scenario, the ELAN response does not reflect any computation relating to the global syntactic structure, and thus it should be completely independent of the well-formedness of the structure. In principle, if participants could be trained to associate particular colors with particular syntactic categories, this account could predict that an ELAN would be observed when the syntactic category prediction initiated be a color was violated. Another prediction of this account is that an increased ELAN should be observed to grammatical sentences if they contain optional modifiers that violate the syntactic category prediction (*Max's <u>very</u> provocative theory*), although one ERP experiment failed to show such an effect (Austin & Phillips, 2004).

On the second account, the ELAN indexes early recognition of ungrammaticality, which is made possible by the syntactic prediction available in the non-ellipsis context.

On this account, in the standard ELAN non-ellipsis context, *Max's of…*, encountering the possessor allows prediction of the upcoming NP. This prediction indicates that until the noun heading the NP is encountered, any incoming material has to be attached within the DP. Therefore, when the parser looks for an attachment site for the word *of*, it need only consider the current 'workspace', consisting of the predicted noun position and potential positions immediately to its left, in order to determine that the sentence is ungrammatical. On the other hand, in the ellipsis context, the possessor may be understood to be followed by a phonetically null noun, allowing the parser to close off the NP, with the consequence that the effective workspace for the attachment of the word *of* becomes the entire sentence structure. In other words, in the ellipsis context, a broader space of alternatives must be evaluated to determine ungrammaticality; the parser needs to check to make sure that the prior context does not provide a position where the *of* phrase can be non-locally attached, as in a phrase such as *the destruction by Max's army of the town.* According to this account, ungrammaticality in sentences like the ellipsis case where there is not a syntactic prediction to narrow the space of possibilities is recognized too late to affect the amplitude of the early part of the ERP; this account has to assume that there is some other explanation for why a 'late' left anterior negativity is not observed to such sentences later in the time-window.

On the third account, the ELAN does not depend on the presence of predictive processes at all, but rather indexes the response to a specific kind of ungrammaticality and not others. On this account, the ELAN indexes ungrammaticality due to the failure to fulfill grammatical requirements of the previous input, but does not index ungrammaticality due to the failure to integrate new input. In the non-ellipsis condition, the preposition cannot be attached inside the current phrase, indicating that it must be part of a new phrase and that the possessor will not fulfill its requirement for a noun. In the ellipsis condition, the ellipsis allows the possessive to fulfill its requirement for a noun, and when the preposition is encountered the ungrammaticality has a different cause, that it simply cannot be attached to any existing phrases in the sentence, and therefore cannot be integrated into the syntactic structure. Therefore, the ELAN may simply be an electrophysiological index of the former kind of ungrammaticality but not the latter (recent work by Kluender and colleagues (Rosenfelt et al., 2009) makes a similar argument, although their manipulation includes additional differences between conditions). On this account, no prediction of the noun need ever have been made or violated, in the sense that no phrase structure need have been built on the basis of the context without having encountered the terminals, and no syntactic category representations need have been pre-activated.

In summary, the results of Experiment 3 demonstrate that the early ERP response to a word that constitutes a phrase structure violation is greater in amplitude when the word prevents a syntactic requirement of the current context from being fulfilled than when the requirements of the context are fulfilled and the word simply cannot be added to the structure. Although both predictive and non-predictive mechanisms can account for these results, one argument in favor of a predictive mechanism might be the earliness of the effect. In the next section I shift to the question of what constraints the timing and localization of such an effect could theoretically put on the timing and architecture assumed for access and integration processes in comprehension.

5.5 Using predictive effects to constrain timing estimates

Predictive effects are often expected to be observed ‘early’ in the processing time course, and conversely, if early effects are observed, they are sometimes explained by appeal to prediction. Thus, in our article reporting the ELAN experiment (Lau et al., 2006), we put strong emphasis on the fact that ELAN effects occurred before the earliest estimates of lexical access in arguing that they must be due to prediction. However, if a predictive effect is observed very early, it not only puts constraints on the timecourse of the predictive mechanism itself, it also puts constraints on the timecourse of the bottom-up processing.

To illustrate this, let’s assume a very simple model with representational levels that are serially ordered on the first feedforward pass through the system (it doesn’t matter if some of the stages overlap, as long as there is some ordering to which levels get the information first). On the first pass, each level does a certain amount of processing before passing information to the next level. Now let’s assume that at some point in each processing stage, a candidate representation is ‘selected’; e.g., once a candidate representation gets 20 units of support, that candidate is chosen and processing moves to the next level. Figure 29 illustrates what the process looks like for a particular word presented in isolation in our toy model.

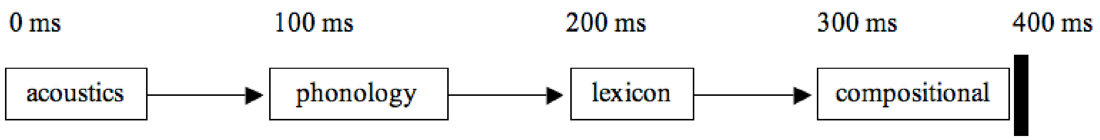


Figure 29. Sample timecourse for feedforward processing of word in context in simplified model.

The earliest point with respect to the input at which predictive pre-activation could facilitate processing is before the input is even encountered, when the predictive information in the context is encountered. At this point, the activation of the predicted representation could be boosted. This could then reduce the amount of processing of the bottom-up input needed for the candidate to reach the threshold (whatever that processing entails). In this case, say the most specific prediction allowed by the context is that the next word is likely to be from the syntactic category 'noun'. Pre-activation based on this prediction would have to be implemented at the lexical level, where the syntactic information associated with each word is stored. The noun category is big enough that, for the moment, I will assume it is unfeasible for any predictions of particular phonology to be maintained at the phonological representation level. Therefore, the first, feed-forward stream of processing will proceed as if there was no supporting context until the lexical representation level, where processing will be speeded, as shown in Figure 30.

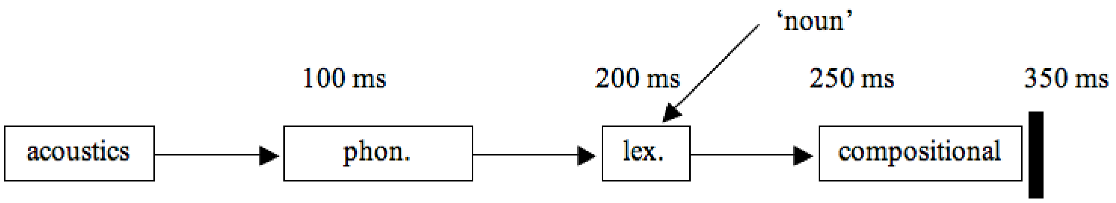


Figure 30. Sample timecourse for processing when the syntactic category of the word can be predictively pre-activated by the context.

What this example illustrates is that one can only find predictive effects that take the form of pre-activation as early as the processing level over which the prediction is made; that is, if the prediction is only as specific as 'noun', then it would be impossible to find effects earlier than the first time at which the earliest feedforward activity reaches the lexical representation level. Similarly, as a semantic category prediction like 'animal' presumably does not entail predictions for particular phonology or orthography, then effects of semantic category prediction should not be observed until the earliest point at which lexical information could be accessed for that word when it was presented in isolation.

If a contextual effect is observed right around the earliest point at which feedforward activity reaches the lexical representation level, this can be taken as an argument that the effect is predictive, if it is assumed that processes involved in selecting a candidate or integrating this candidate with the previous sentence material take some amount of time. According to this assumption, if the effect reflected a response to some aspect of the relationship between the critical word and the previous context, it would only appear with some delay after the point at which the earliest feedforward activity reached the critical word (in this example in Figure 31, not until 250 ms).

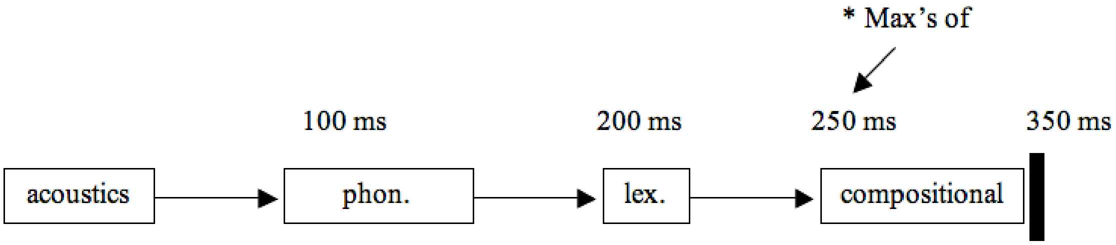


Figure 31. Sample timecourse for processing for a non-predictive effect of syntactic context, in which differences in context do not affect processing until the information from the bottom-up input is actually combined with the previous context to update the larger representations of the sentence (e.g. syntactic structure) being constructed.

In some cases, a predictive effect may appear to onset earlier than the first point at which feedforward activity reaches the lexical representation level. The ELAN is such an example, as its normal onset (150-200 ms) is slightly earlier than the 200 ms range at which some previous authors have placed lexical access (Allopenna, Magnuson, & Tanenhaus, 1998; van Petten et al., 1999). If we continue to assume a pre-activation mechanism, there are two ways out of this apparent paradox. The first is to claim that the earliest contact with the lexical level of representation is earlier than had been originally assumed. For example, it might be the case that, as Friederici (2002) suggested, syntactic category information is accessed earlier than most other kinds of information linked to the lexical representation; also, most demonstrations of the ELAN have involved targets that were function words or had functional affixes, which might be accessed faster than the average content word. The first contact with lexical representations in general could also just be much earlier than previously assumed, as recent authors have argued (e.g., Penolazzi, Hauk, & Pulvermuller, 2007). The second option is to claim that the content of the prediction is actually relevant for an earlier level of processing. Dikker and colleagues (Dikker et al., 2009) suggest that 'predictions about upcoming word categories include form-based estimates', in other words, that prediction of a noun actually initiates prediction of specific visual forms in visual cortex or the visual word form area that are distinct from the visual or orthographic features present in the function words they used, assuming the view that different syntactic categories have different phonological tendencies (Farmer et al., 2006). In more recent work, Dikker and colleagues provide evidence than phonologically typical words are more effective at eliciting early prediction violation responses, supporting this explanation (Dikker et al., submitted).

An alternative means of explaining very early effects through prediction is to assume a different predictive mechanism than pre-activation, which can make it possible for a context that only makes specific predictions about the lexical level to have effects earlier than the first point at which bottom-up input makes contact with the lexical level. This could happen if the system can optimize the duration of the processing stages based on the *amount* of information available rather than the particular kind of information available. This would be something like the system saying, ‘I know going into this that recognition is going to be easier than in isolation, because I have information that is going to let me winnow down the candidates at the lexical representation level. Therefore, I can skip some of the computations that I would normally do at the phonological level, because the extra information I am going to get at the lexical level will probably give me enough evidence to decide among the candidates anyway.’ In this system, the amount of uncertainty at one level would therefore gate the amount of effort spent in preprocessing at an earlier level, perhaps by lowering the threshold for selecting candidate representations or perhaps by moving on before a single candidate has been selected. This would allow information to reach the lexical level faster than normal. In the toy model we have been considering, this would mean that following a strongly predictive context (which would reduce uncertainty), phonological processing would be reduced, lexical processing would begin earlier, and thus violations of lexical-level predictions like syntactic category could be observed as early as 150 ms in this example (Figure 32).

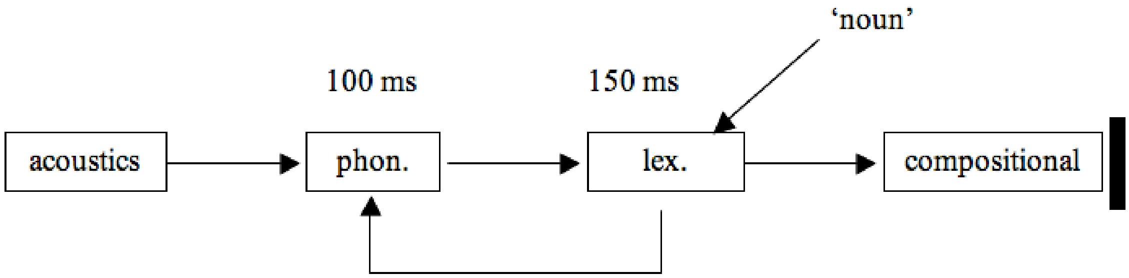


Figure 32. Sample timecourse for processing in a case in which the amount of prior information about the upcoming input actually alters the dynamics of the first feedforward flow of information, by reducing the amount of computation at earlier stages if information contributing to identification is already available at more abstract representational levels.

To sum up, if we assume that the timing of feedforward stages is fixed, and if we observe contextual effects earlier than the earliest point at which the level of information constrained by the context begins to be accessed through feedforward activity, this indicates that something is wrong with our hypotheses about when those feedforward stages occur or about what level of information can be predicted by the context. If we observe contextual effects exactly around the earliest point at which the level of interest is accessed through feedforward activity, and if we assume that selection and composition of representations requires processing time, we can argue that the contextual effects reflect prediction. If we assume that the timing of feedforward stages can change dynamically, all bets are off.

Currently, the timing with which various levels of representation are accessed during language comprehension is a huge topic of debate (see Almeida, 2009, for discussion). Therefore, it is difficult to use the timing of feedforward stages to constrain the interpretation of contextual effects. However, as we achieve more consensus on the feedforward timecourse is achieved, this possibility should become more available; in the meantime, very early contextual effects such as the ELAN may themselves provide some constraint on the feedforward timecourse, as illustrated above.

5.7 Conclusion

In this chapter I presented an ERP experiment showing that phrase structure violations that are due to the failure to fulfill a syntactic requirement of the prior context demonstrate a very early ERP response that is not observed for other kinds of syntactic violations. Although several interpretations of this result are possible, the data are consistent with the hypothesis that phrase structure requirements are instantiated as syntactic predictions in processing, either through pre-activating words of a particular syntactic category or through pre-constructing upcoming syntactic positions. I suggest that current estimates of the timecourse of feedforward processes in lexical access may be too contentious to justify taking the earliness of the effect as evidence for a particular underlying mechanism, but the earliness of the effect may put some outer constraints on those estimates under certain architectural assumptions.

Chapter 6: Syntactic prediction II – Non-adjacent dependencies

6.1 Introduction

In the preceding chapters, I have examined empirical means of showing evidence for prediction during language comprehension, and I have proposed a neural model of how lexical and conceptual predictions may be implemented through top-down pre-activation. The implicit assumption is that top-down activation of representations is beneficial because it both reduces the need for bottom-up computation and it increases the likelihood that the correct representation will be activated when the bottom-up input is noisy or ambiguous.

In this chapter, I turn to a more complex and interesting way in which predictions can make comprehension more accurate. I discuss a number of behavioral findings in the processing of syntactic dependencies—relationships between different parts of the sentence governed by the grammar—that suggest that predictions can make comprehension more accurate by obviating the need for subsequent memory retrieval, which may be prone to error.

One way of checking that a syntactic dependency between two non-adjacent elements is grammatically fulfilled would be to wait until the second element is encountered and then retrieve the necessary information from the first element from memory. Alternatively, if the first element unambiguously indicates the beginning of a dependency, it could trigger the pre-construction of the structural position of the second element of the dependency, with properties required of the second element included as part of the prediction. In this way, the dependency chain between the first and the second element of the dependency could be built in advance, and simply await confirmation that the bottom-up evidence provides a candidate with the required properties in the appropriate position. This kind of prediction would make it unnecessary to wait until the second element is encountered to retrieve the first element from memory to license it, an operation that could be prone to interference from other items in memory.

I will argue that a pattern of selective fallibility observed in dependency processing—immediately accurate analyses under some conditions and erroneous analyses under other conditions—supports the claim that when this kind of prediction is available in comprehension, it makes processing of syntactic dependencies more accurate. I will present behavioral evidence from the processing of non-adjacent subject-verb agreement, anaphora, and other kinds of dependencies that suggests that errors only occur when retrieval is necessary, and that potential errors in analysis are avoided when prediction is available. These generalizations rest on the contributions of a number of researchers, but especially Matt Wagers and Colin Phillips, and receive fuller presentation in Phillips, Wagers, and Lau (submitted).

6.2 Advantages of prediction: predicting features

To provide an account for how prediction can make comprehension more accurate, we need to start with a case in which performance is not so accurate, and we need to understand why it is not accurate. One such case that I have examined in detail with Matt Wagers is the phenomenon of agreement attraction.

In English, subjects and verbs must agree in number. However, agreement in language production is famously prone to a particular kind of error, known as agreement attraction, in which the verb fails to match the agreement features of its grammatical controller (the head noun of the subject phrase) and instead takes on the features of a nearby but non-controlling ‘attractor’ (18).

(18) The sheer <u>weight</u> of all these figures <u>make</u> them harder to understand.

These kinds of errors are relatively common, elicited about 13% of the time in production experiments when participants are provided with appropriate preambles, and observed frequently in natural speech and in texts from newspapers to academic journals. An analogue to the production effects can also be observed in comprehension, in the form of reduced disruption to ungrammatical subject-verb agreement when agreement attraction is possible (Pearlmutter, Garnsey, & Bock, 1999; Clifton, Frazier & Deevy, 1999; Kaan, 2002; Häussler & Bader, 2007); indeed, these sentences intuitively sound better than their non-attraction counterparts (*The key to the cabinet were on the table* vs. *The key to the cabinets were on the table*).

Production studies have established a number of factors that govern the production of such errors, including which number features are involved, the structural depth of the attractor with respect to the grammatical controller, and linear order (Bock & Miller, 1991; Bock & Cutting, 1992; Bock & Eberhard, 1993; Vigliocco & Nicol, 1998; Hartsuiker, Antón-Méndez, & Van Zee, 2001; Haskell & MacDonald, 2005; Thornton & MacDonald, 2003). One of the most important generalizations across these studies is that attraction effects are morphologically selective. In English, robust agreement attraction only occurs when the subject is singular and the verb is plural (*The key to the cabinets were on the table*). If the subject is plural and the verb is singular (*The keys to the* *cabinet…*), very few agreement errors are observed (~3% according to Eberhard, Cutting, & Bock, 2005). This pattern is usually chalked up to some consequence of the plural being 'marked' (e.g. Eberhard, 1997; Kimball & Aissen, 1971).

The other relevant generalization from the production literature is that the occurrence of errors does not seem to depend on linear proximity of the attractor noun to the verb, as is often suggested (e.g. Quirk, Greenbaum, Leech, & Svartvik, 1985). One might have thought, following models of comprehension in which surface statistics play a prominent role (e.g. Tabor, Galantucci, & Richardson, 2004), that it is the fact that plural nouns are often immediately followed by plural verbs that causes errors like *The key to [the cabinets were]*. However, several lines of evidence argue against this interpretation. First, on this account, the morphological selectivity of the effect is unexpected; since singular nouns are usually followed by singular verbs, this account would predict that singular attraction errors like *The keys to the cabinet was* should sound equally good and occur equally frequently, which they do not. Second, a number of production studies show that linear adjacency of the attractor and verb, even when the attractor is plural, does not necessarily lead to strong attraction (Bock & Cutting, 1992; Solomon & Pearlmutter, 2004), and that attraction can occur without linear adjacency, being reliably elicited in configurations such as *Are the key to the cabinets on the table?* (Vigliocco & Nicol, 1998).

The most influential theory of agreement attraction in the production literature argues that attraction is a result of feature movement or 'percolation' within a syntactic representation (Nicol, Forster, Veres, 1997; Vigliocco & Nicol, 1998; Franck, Vigliocco & Nicol, 2002; Eberhard, Cutting & Bock, 2005). On this account, information is sometimes spuriously transmitted through the structural links between constituents. In percolation, features on a given syntactic constituent can be transferred to other, nearby constituents, but they can only be transferred one syntactic ‘step’ at a time; in other words, features must pass first to the immediately dominating syntactic node, then to the next, and so on. This ‘stepwise’ movement is reflected in the reduced likelihood of feature movement with increasing syntactic distance between nodes. In the typical agreement attraction case of a subject with a PP modifier (*the key to the cabinets*), the number features bound to the noun phrase within the PP percolate upward, valuing higher phrasal projections for number. In some proportion of cases, these features can erroneously percolate up to the highest projection, that of the subject noun phrase. By hypothesis, the verb or verb phrase is reliably valued by the number on the subject phrase, and so will be inappropriately valued in just that proportion of cases when the PP-object’s number percolates far enough to value the subject phrase. The percolation hypothesis is supported by evidence showing that for subjects with two prepositional modifiers (N-PP1-PP2), the PP modifier that is structurally closer to the subject head (and incidentally, further from the verb; see Figure 33) induces more attraction (errors PP1 > errors PP2; Franck, Vigliocco, & Nicol, 2002) and that local nouns in a PP modifier configuration induce more attraction than local nouns embedded in a relative clause, which are structurally more distant from the subject head.

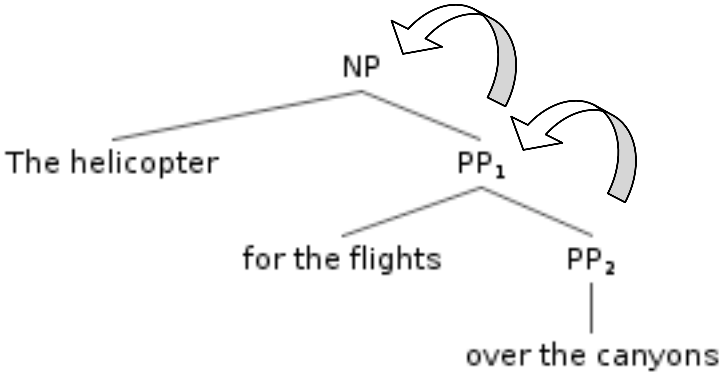


Figure 33. Illustration of how a plural number feature could 'percolate' up the structural tree. Note that the plural feature on *flights* would just require fewer movements than the plural feature on *canyons* to erroneously mark the subject as plural.

However, based on a series of experiments (Wagers, Lau, & Phillips, 2009), we have argued that, at least in comprehension, percolation is unlikely to account for agreement attraction errors. First, we demonstrated effects of attraction in the configuration shown in (19), in which the attractor (the plural relative clause head) does not intervene, either linearly or hierarchically, between the subject and the verb. Attraction in this configuration is hard to capture on a traditional percolation model, because it would require both upwards and downwards percolation for the plural features on the relative clause head to end up on the relative clause subject (Figure 34).

(19) *The <u>musicians</u> who the reviewer <u>praise</u>…

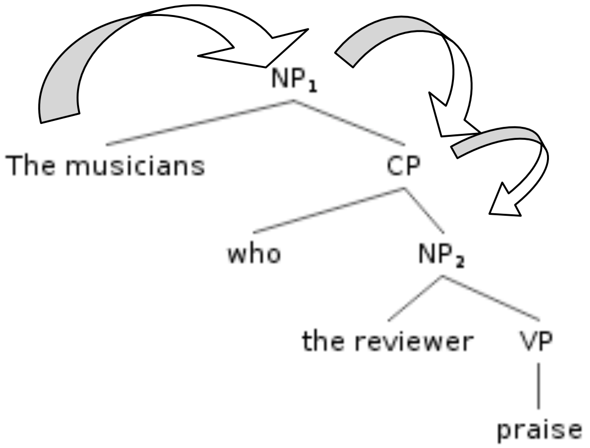
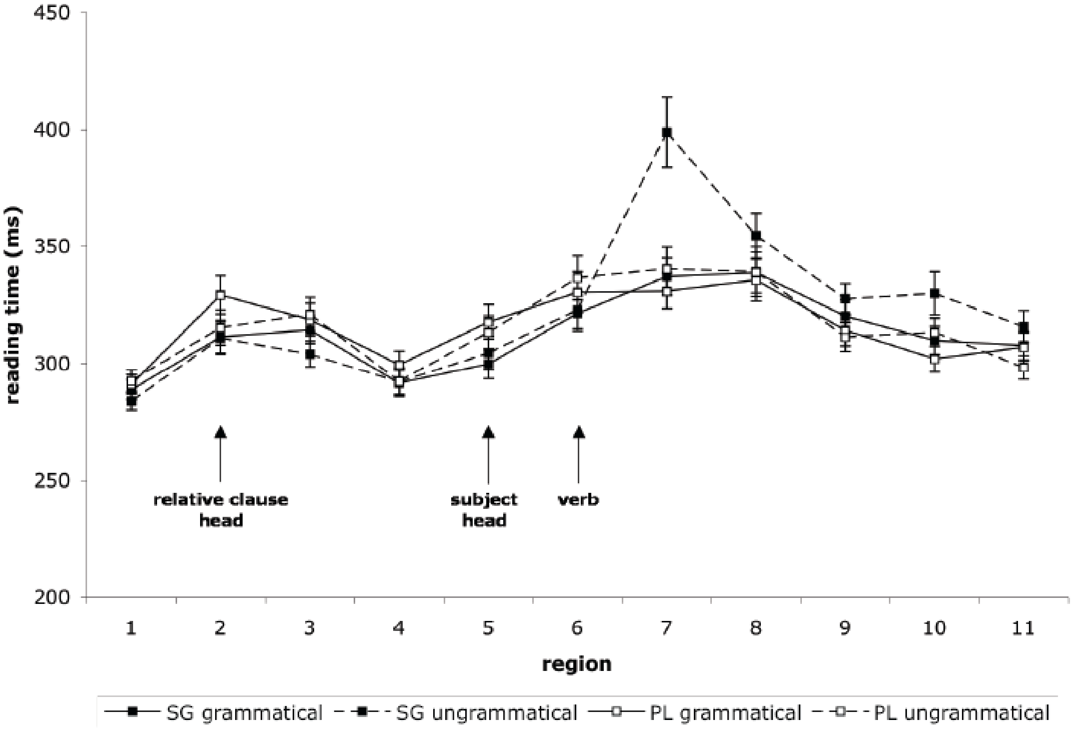


Figure 34. Illustration of the up-and-down percolation path required to capture attraction in (19).

We showed a classic profile of agreement attraction in reading times: a large slow-down, relative to grammatical controls, for singular subject-plural verb agreement errors in the absence of a plural attractor, but almost no measurable slow-down for the same errors when a plural attractor was present[7] (Figure 35).



The$_1$ musician(s)$_2$ who$_3$ the$_4$ reviewer$_5$ praise(s)$_6$ so$_7$ highly$_8$ will$_9$ probably$_{10}$ win(s)$_{11}$ ...

Figure 35. Self-paced reading results from Experiment 2 of Wagers, Lau, and Phillips, 2009. Region by region means segregated by relative clause head number and grammaticality. Error bars indicate standard error of the mean.

Second, across six attraction experiments we failed to show evidence of a key prediction of percolation theories, that erroneous percolation should disrupt grammatical

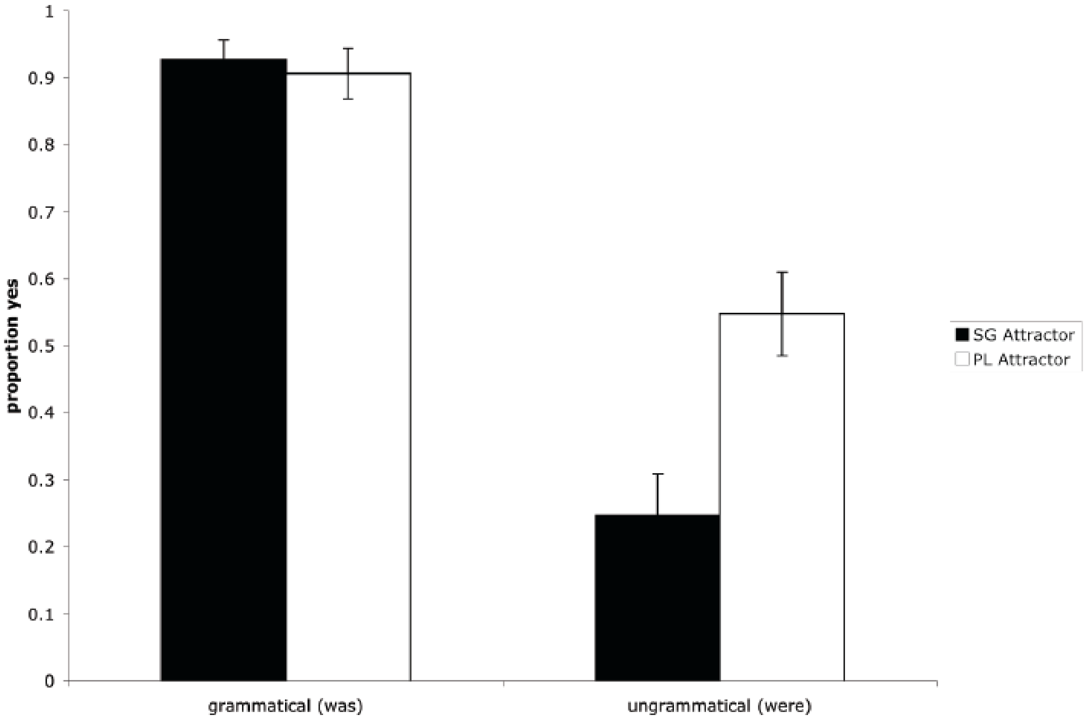
---

[7] Note that in a replication of this experiment (Experiment 3 of Wagers, Lau, & Phillips, 2009) we observed the same significant effect of attraction, but in this case it did not completely eradicate the disruption due to ungrammaticality.

sentences to the same degree that it disrupts recognition of ungrammaticality. In other words, since percolation depends only on features of non-head nouns percolating over to mark the subject, this should happen whether or not the verb happens to match the attractor in number. Therefore, if on 20% of ungrammatical trials, plural features percolate up to the subject and facilitate reading times on the plural verb, then on 20% of grammatical trials, plural features should percolate up to the subject and slow down reading times on the grammatical singular verb. However, as the matched solid lines in Figure 35 illustrate, we observed no difference in reading times for the grammatical sentences. We replicated this observation in another experiment with the same structures and in a number of other self-paced-reading experiments using the complex subject *key-to-the-cabinets* configuration (modulo an additional plural complexity effect; see Wagers et al., 2009). We have found the same pattern in a number of studies in speeded judgments to sentences presented with RSVP: reduced accuracy in judging ungrammatical attraction sentences as incorrect, but no effect on accuracy in judging grammatical attraction sentences.

Based on these results, we have argued instead, following other authors (Badecker & Lewis, 2007) that agreement attraction errors in comprehension are due to failures in a content-addressable retrieval mechanism used to check agreement routinely during sentence comprehension. The fact that only sentences in which subject and verb number mismatched showed a reliable attractor effect suggests that information from the verb (or auxiliary) plays a necessary role in attraction effects online. A content-addressable-retrieval mechanism that uses the information on the verb could naturally give rise to the observed pattern of attraction (Gillund & Shiffrin, 1984; Lewis & Vasishth, 2005;

McElree, 2006). The core idea behind such a retrieval mechanism is that features in the current input are used as cues to query the contents of memory simultaneously, much like keywords are used in Internet search engines to find matching websites. The results of such a query could match all keywords or only a subset of them, and this degree of match may affect the likelihood of retrieving the results from memory. A content-based retrieval mechanism can give rise to errors through retrieval of partially-matching but erroneous items (e.g., Gardiner, Craik & Birtwistle, 1972).



The key to the cabinets(s) was/were rusty from many years of disuse.

Figure 36. Speeded acceptability judgment results from Experiment 7 of Wagers, Lau, & Phillips (2009). Mean proportion 'acceptable' responses by grammaticality and attractor number. Error bars indicate standard error of the mean.

In attraction, numerous sources of information provided by the verb may form the cues for retrieval. For the configurations we have considered, we assume that the retrieval cues consist of (privatively specified) agreement features, like [Number:Pl], structural cues, like [Case:Nom] or [Role:Subj] that identify the subject, and clause-bounding cues. When neither of the NPs matches the combined cue, as in the ungrammatical sentences, the number-matching non-subject is sometimes the best match. This can account for the errors observed[8] (see Wagers, 2008, for more discussion).

Having argued that similarity based interference in cue-based retrieval is the source of errors in agreement attraction, I now want to return to the original question: What role does prediction play in avoiding such errors and making comprehension more accurate? The key idea is that content-based retrieval is inherently prone to error and that prediction makes retrieval unnecessary. As I discuss above, erroneous retrieval of partial matches is always a danger for content-based retrieval mechanisms. However, if the information about subject number can be realized as a prediction for verb number, the bottom-up verb input can simply be checked against the prediction, without a retrieval operation. This can account for the lack of errors in the grammatical cases.

What this view requires is that a fulfilled prediction for verb number either is equivalent to or satisfies an agreement-checking operation. On one implementation, agreement checking just *is* satisfying a prediction of number. When an agreement-marked element is encountered, a prediction is made for the number-marking at other structural positions. When the number of both elements matches, the prediction is satisfied, and nothing else happens. When the number of the verb does not match the subject, the prediction is not satisfied, and this triggers a re-assessment of the previous material. We suggest that this reassessment may take the form of a content-based retrieval partially on the basis of the number cues at the verb to find an element that would license the number,

---

[8] Although one problem that may arise with scaling such a model up to account for comprehension across a wider set of data in English is the limited amount of agreement information that is available on most verbs in English. See Solomon and Pearlmutter (submitted).

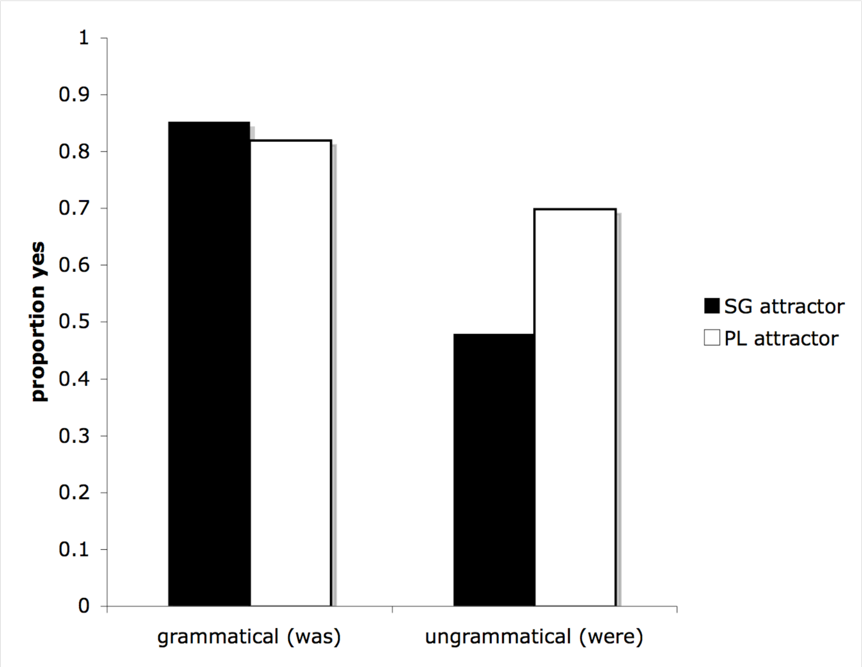
or in other words, justify the presence of plural number on the verb. This is why, in ungrammatical attraction cases, we observe what we interpret as evidence of erroneous retrieval. On a subtly different variant of this approach, a separate agreement-checking operation occurs, and this operation is fulfilled predictively when the number of the verb can be predicted. If the input does not fulfill the number prediction, all the predicted operations are cancelled and must be re-done, including agreement checking. Prediction of morphosyntactic features may also facilitate morphological decomposition, by indicating whether or not the upcoming stimulus will need to be parsed into different morphemes (Lau, Rozanova & Phillips, 2007).

Although prediction provides a nice account for the pattern of results observed in agreement attraction, a non-predictive account is also possible. On one such account, subject-verb agreement always requires a retrieval-based checking operation using cues such as agreement features like [Number:Pl], structural cues like [Case:Nom], and clause-bounding cues. The selective pattern of errors arises, under this account, if the mechanism is set up so that it almost never retrieves a partially matching item when a fully matching item is available. This is a property of standard content-addressable memory models when cue combination rules are supralinear (e.g. Gillund & Shiffrin, 1984). In the grammatical cases, a full match would always be available, while in the ungrammatical cases, only partial matches would be available, and sometimes a non-head item would be selected erroneously based on the partial match, allowing the agreement checking operation to falsely come to the conclusion that the verb number was licensed. However, the prediction-reanalysis view has several advantages over this pure retrieval account. Given that English agreement paradigms for lexical verbs are largely syncretic, it may be necessary to use top-down information, like the number of the subject head, to identify the number features of the verb in the first place. Also, McElree, Foraker, & Dyer (2003) have argued on the basis of SAT time-course dynamics that adjacent subjects and verbs are integrated without necessitating a memory retrieval. However we observe in the relative clause attraction cases that adjacent subjects and verbs are nonetheless liable to attraction when subject and verb mismatch (*the musician who the reviewer praise*). These two findings are naturally reconciled if no retrieval is required for matching subjects and verbs, but a reanalysis-cued retrieval is required for subject-verb mismatches.

We recently attempted to adduce additional evidence that prediction is the critical factor in making agreement processing accurate with a variant on our original attraction design. We constructed sentences that were designed not to predict an upcoming verb at all, with the hypothesis that this should lead to increased errors even in the grammatical case, if it is the case that avoidance of attraction is a consequence of predictive processes. We accomplished this by adding a coordinated predicate phrase, which is structurally optional (20). According to our hypothesis, the singular subject predicts a singular verb, which is satisfied by the modal 'could'. However, there is no reason for the singular subject to predict a subsequent singular verb in a coordinate phrase. Therefore, when the second verb is encountered, cue-based retrieval should be necessary to make sure that the number on the verb is licensed.

(20)
a. The slogan on the poster couldn't be read easily but was designed to get attention.
b. The slogan on the posters couldn't be read easily but was designed to...
c. The slogan on the poster couldn't be read easily but were designed to...
d. The slogan on the posters couldn't be read easily but were designed to…

We tested the four conditions in (20) with a speeded acceptability judgment paradigm with RSVP in a pilot study with 16 participants. In contrast to our prediction, the results were the same as for our original attraction items; significantly more errors in the ungrammatical attraction case, but not in the grammatical attraction case (Figure 37).



The slogan on the poster couldn't be read easily but was designed to get attention.

Figure 37. Speeded acceptability judgment results from a pilot study investigating attraction in a coordination structure in which the agreeing verb could not be predicted. Mean proportion 'acceptable' responses by grammaticality and attractor number.

Although these results fail to provide decisive evidence in favor of the prediction hypothesis, they also cannot be taken as conclusive evidence against it. First, overall accuracy for these materials in this group of participants was markedly worse than in our previous experiments; for example, nearly 50% of the ungrammatical controls were incorrectly labeled as acceptable, while this number was closer to 20% in the one clause experiment above. This may have been partially due to the filler items, which included a subset of fairly complex sentences that may have caused some participants to get tired or lose focus. Second, the significant number of coordinate clauses presented in the course of the experiment may have led to the prediction of the second verb in some proportion of cases, which would preclude the need for retrieval in the grammatical attraction cases. These potential confounds could be resolved in future experiments.

On the other hand, a recent SAT (speed-accuracy tradeoff) experiment provides new evidence for prediction in agreement processing but suggests that perhaps only morphologically explicit features can be predicted (Wagers, Lau, Stroud, McElree, Phillips, 2009). This experiment contrasted singular and plural determiner-noun number mismatch in sentences like (21). The assumption is that when the determiner and noun are adjacent, the features of the determiner should still be available at the noun and no retrieval should be necessary. However, when the determiner and noun are separated by intervening adjectives, retrieval of the determiner should be necessary to check that it licenses the noun number, unless prediction makes it unnecessary.

(21) Yolanda always remembered
 this {mischievous, face-making} <u>monkey</u>.
*this {mischievous, face-making} <u>monkeys</u>.
 these {mischievous, face-making} <u>monkeys</u>.
*these {mischievous, face-making} <u>monkey</u>.

The results suggest that memory retrieval is necessary to recognize the number mismatch for singular determiner – plural noun combinations (*this monkeys*) across intervening material, as accuracy at detecting the violation decreased with increasing distance. However, no such effect of distance on accuracy was observed for detecting number mismatch in plural determiner – singular noun combinations (*these monkey*), arguably because the plural determiner leads to a prediction for a plural noun, negating the need for retrieval to check agreement at the noun[9].

If this interpretation of the data is correct, the story above would need to be modified somewhat, because it suggests that prediction only obviates the need for retrieval in some cases—i.e. when the number on the first element of the dependency is the marked case or has overt morphological consequences for the second element. An ERP study currently in progress is aimed at providing more evidence on whether morphosyntactic prediction is in fact limited in this way; preliminary results suggest distinct response patterns for the two kinds of determiner-noun mismatches.

To sum up the results of this section, I have demonstrated a pattern of selective fallibility in the comprehension of subject-verb agreement: errors are selectively observed in ungrammatical sentences. I have argued that the errors that are observed are due to erroneous content-based retrieval of partial matches, and I have argued that prediction may be the mechanism that prevents errors from being observed in the comprehension of grammatical sentences. In the next section, I discuss several other cases of selective fallibility that may be accounted for through predictive mechanisms.

6.3 Advantages of prediction: predicting structure

Another well-studied class of dependencies are those in which a pronoun or anaphor must be co-indexed or co-referenced with another referent in the sentence or discourse. In this section, I will describe a pattern of selective fallibility in pronoun

---

[9] One potential confound in these materials is that the plural determiner-singular noun sequence (*these monkey*) could be grammatically continued if the noun serves a modifier role (*these monkey catchers*) (Kennison, 2005), while the singular determiner-plural noun sequence (*this monkeys*) could not. However, because a period indicated that the sentence ended with the noun, it seems unlikely that participants considered this possibility.

processing that seems to potentially lend itself to an account like the one we proposed for agreement, in which prediction of the dependency allowed the parser to avoid an error-prone retrieval process. Pronouns can appear in one of two configurations: *forwards anaphora*, in which the referent linearly precedes the pronoun 0 and *backwards anaphora*, in which the pronoun linearly precedes the antecedent (23). Note that in the backwards anaphora case, the referential dependency is signaled at the first element of the dependency, by the pronoun, while in the forwards anaphora case, the referential dependency is not signaled until the second element of the dependency is encountered. As I will describe below, the accuracy with which the comprehender initially processes the dependency within the constraints of the grammar seems to depend on the degree to which the comprehender is able to predict the upcoming dependency chain.

(22) John$_i$ said he$_i$ will go to the store later.
(23) While he$_i$ was on the phone, John$_i$ put the pie in the oven.

Binding Principle B of classic binding theory (Chomsky 1981) blocks referential interpretations in which the pronoun is co-indexed with an antecedent within the same clause, accounting for the unacceptability of sentences like 0. A question of significant interest has been how such interpretations are ruled out during online processing: are interpretations unacceptable according to this binding constraint ever erroneously considered?

(24) *John$_i$ likes him$_i$.

Results on this question have been mixed, but at least some evidence suggests that erroneous interpretations are temporarily constructed. Typically, online processing of anaphora is examined by manipulating the match of various potential antecedents to the pronoun on features like number or gender, on the assumption that the match between pronoun and referent will only affect processing if the referent is being considered as a potential antecedent. Although several studies have failed to show effects of referents that would not be allowed as antecedents according to Principle B (Nicol & Swinney, 1989; Clifton, Kennison, & Albrecht, 1997; Lee & Williams, 2006), other studies (Badecker and Straub, 2002; Kennison, 2003) do report such effects. For example, Badecker and Straub found that reading times at the pronoun in (25) were affected by the gender of the local subject, even though this subject was not an acceptable antecedent. Results like this have led such researchers to the conclusion that, at least in some cases, comprehenders temporarily consider elements in unacceptable positions as potential antecedents.

(25) John$_i$ said that Bill/Beth owed <u>him</u>$_i$ another chance to solve the problem.

Backwards anaphora is governed by another binding constraint, Principle C. Principle C rules out coreference between a pronoun and a referent that c-commands it, as in (26).

(26) *He$_i$ put the pie in the oven while John$_i$ was on the phone.

Backwards anaphora has been shown to be processed 'actively': the presence of a pronoun at the beginning of the sentence seems to drive an active search for the referent. Therefore, sentences like (27) demonstrate a slowdown in reaction times compared to (28) at the noun (van Gompel & Liversedge, 2003), even though it would be perfectly acceptable to have the referent for the pronoun appear somewhere else in the sentence, as in (29). This is sometimes referred to as the *gender mismatch effect*.

(27) While he was on the phone, John put the pie in the oven.
(28) While he was on the phone, Mary put the pie in the oven.

(29) While he_i was on the phone, Mary put the pie in the oven for $John_i$.

The order in which information becomes available in the backwards anaphora configuration is different than in the forwards anaphora configuration; whereas in the forwards anaphora case, the search for a referent operates over preceding material, in the backwards anaphora case in which there is no prior context, the referent must be found in subsequent input. This raises the question of whether nouns that are in unavailable positions may still be temporarily considered as referents, as seems to be the case for forwards anaphora.

To determine whether Principle C constrains online processing, we tested whether the gender mismatch effect described above would be observed for referents that would be ruled out by Principle C, as in (30), where the potential referent is c-commanded by the pronoun (Kazanina, Lau, Lieberman, Phillips, & Yoshida, 2007). We replicated the gender mismatch effect for referents in licit positions (31), but found no evidence of the gender mismatch effect for referents in illicit positions[10], across three experiments that tested three different configurations that are subject to Principle C.

(30) Because yesterday he was at a party while John/Mary was making a pie, Mike…
(31) Because yesterday while he was at a party John/Mary was making a pie, Mike…

---

[10] All sentences in the experiments ultimately contained an available referent for the pronoun, as in the examples here. We also included a control condition without a pronoun *(Because yesterday while John was at a party Mary…)* in order to ensure that mismatch effects were not due to the cost of introducing a new referent into the discourse. See Kazanina et al. (2007) for more detail.

Overall, then, we see a contrast between the extent to which grammatical constraints are accurately considered in the first stages of processing referential dependencies involving pronouns: the processing of backwards anaphora seems to immediately respect Principle C, while the processing of forwards anaphora seems to sometimes erroneously consider interpretations that would violate Principle B. The question I would like to consider now is whether this contrast can be explained in the same way that we explained the contrast between agreement attraction in grammatical and ungrammatical sentences, through a contrast between predictive processing and retrieval.

In forwards anaphora cases, there is no indication that a dependency must be formed until the second element of the dependency, the pronoun, is encountered. This means that when the pronoun is encountered, a retrieval process is required to find the antecedent. If we assume that the mechanism is cue-based retrieval, and that cues such as number, gender, and especially structural position/clause membership are used to find the appropriate referent, we can account for the effects of illicit antecedents observed through erroneous retrieval of items that partially match the search cues. This part of the argument is straightforward.

However, for prediction to account for the lack of effects of illicit antecedents in the backwards anaphora case, it must explain why incoming nouns in illicit positions do not interfere in an analogous way during forwards search for the antecedent. The backwards anaphora case is different than other cases of syntactic prediction that I have discussed, such as the prediction of a determiner for a noun with matching features or the prediction of a subject for a verb with matching features. In those cases, the predictions involve the content of an item that is expected to occur in a single well-defined structural position, which makes it easy for the system to assess whether that expectation is met or not. In the backwards anaphora case, it is not clear whether the processor is able to form such a definitive expectation about where the antecedent will occur. In the example in (31), one might argue that because the subject position of the main clause could have been predicted on the basis of the subordinate conjunction *while*, this position is then available to be predictively marked as referentially linked to the pronoun, and that the same thing happens with the outer subordinate conjunction *because* in the Principle C example in (30). However, in other examples that show gender mismatch effects, such as (32) (Kazanina et al., 2007, Experiment 3), it is not so obvious that the position that shows the mismatch effects could have been predicted when the pronoun was encountered. Note that this position did not show mismatch effects in a Principle C configuration when it was c-commanded by the pronoun (33).

(32) His/her managers chatted with the fans while the talented young quarterback…
(33) He chatted with the fans while the talented young quarterback…

In order to capture the mismatch effects in this second configuration, we must add something to the account. One possibility is that a specific structural prediction is constantly maintained once the pronoun is encountered, but that it is continually being updated with little cost as more input is encountered. The main assumption is that at each point the system predicts the minimal amount of additional phrase structure necessary to support the predicted referent. For example, in (32), the referent for the pronoun might initially be predicted as an object to the upcoming verb. When the verb *chatted*, whose subcategorization frame does not allow an object, is encountered, the initially predicted position is abandoned. What is tricky for this account, which needs to say that a prediction is always maintained, is where the referent can now be predicted, since there is nothing in the current input to indicate what positions might be made available in the future. The referent could be predicted into a position in some modifier of the verb, which could be abandoned and re-predicted when the preposition *with* is encountered if the prediction was not for a PP modifier, and then abandoned again and re-predicted when *fans* is encountered, which mismatches the number of the pronoun[11]. This process would go on until the correct referent is encountered in the *while* clause. In the Principle C configuration (33), the referent for the pronoun would initially be predicted into the subject position of the next non-subordinate clause, which is the minimal amount of phrase structure that will allow a grammatical dependency to be constructed. This will allow all the intervening material until the next non-subordinate clause to be ignored for the purposes of computing the dependency.

Although this process of frequent prediction and abandonment of prediction seems to give the correct results in this example, more investigation would be needed to see if this algorithm would work in all cases. Another possibility is that the expectation at the pronoun in the backwards anaphora case takes a very abstract form akin to 'I expect to form a referential dependency with the next referent that I encounter that is outside of my c-command domain', rather than a prediction about a particular part of the structure. Why is it the case then that when the parser searches backward through the structure for a referent, it makes mistakes, but when it searches forward through the structure for a referent, it doesn't? Here one could appeal to some other difference between forwards

---

[11] Note that this account would predict a mismatch effect on *fans* in both conditions of (32) relative to both conditions of (33).

and backward searches. For example, in forwards search the candidates can be evaluated serially, while the cue-based retrieval mechanism assumed for backwards search requires parallel access. Another difference that one could appeal to is that in forwards search, the candidate position for the second element of the dependency can usually be evaluated at least a little bit in advance of the bottom-up input associated with the lexical material itself, which is not the case in backwards search. However, more work would be needed to explain exactly why these differences would lead to more rapid fidelity to the grammar in forwards search.

To summarize the discussion so far, we have examined the sensitivity of the processing of referential dependencies to grammatical constraints in two configurations, one in which the order of the dependent elements allows an expectation for the second element to be triggered by the first element (backwards anaphora), and one in which the need to form a dependency is only recognized at the second element (forwards anaphora). We observed that in the processing of forwards anaphora, whose ordering would seem to require a retrieval mechanism, constraints governing dependency formation appear to be initially violated online. This fits the generalization that the processing of dependencies suffers in accuracy when retrieval is necessary. However, although we observe that the ordering of elements in backwards anaphora may allow the referential dependency to be computed predictively, it is not yet clear what it is about the predictive mechanism that allows the parser to avoid attempting to bind structurally unavailable antecedents; I have discussed several possible explanations.

One further caveat to this account is that not all kinds of forwards anaphora are prone to error in considering grammatical constraints in online processing. Reflexives are constrained in their choice of antecedents by Principle A, which essentially forces the antecedent to be the subject of the same clause (34). Evidence from reaction times, eyetracking, and ERPs suggests that this constraint is respected from the earliest points in processing (Nicol & Swinney, 1989; Sturt, 2003; Xiang, Dillon & Phillips, 2009), even though the dependency is not predictable based on its left-hand element. Assuming that predictability is what differentiates the accuracy of respecting Principles B and C online, another account must be given for the accuracy of respecting Principle A, perhaps through its locality (see Wagers, 2008; Phillips et al., 2009).

(34)
a. John$_i$ liked himself$_i$
b. The man$_i$ that the Mary worked for liked himself$_i$.
c. *John$_i$ hoped that Mary liked himself$_i$.

On the other hand, one piece of evidence in favor of this account is that another type of dependency whose first element makes available a similarly vague structural prediction, wh-dependency, also respects grammatical constraints during processing. In wh-dependencies, otherwise known as filler-gap dependencies, a word or phrase is displaced to another part of the sentence, as in wh-questions or relative clauses (35).

(35)
a. What did the teacher assign __?
b. What does the teacher think the children expect her to assign __?
c. The homework that the teacher assigned __ was difficult.

The 'head' of this dependency is the displaced element. At either the wh-element or at the complementizer or second subject that indexes the beginning of the relative clause, it becomes clear to the comprehender that an element has been displaced. The 'foot' of this dependency is considered to be the gap where the element came from and where there is now an argument missing. However, the only way of determining for certain that there is a gap is to wait until the following word is encountered, which will show that the argument is missing. Then, to ensure that the sentence is appropriately interpreted and considered well-formed, some kind of operation must be done to link the displaced element with its original position. I will refer to this as 'constructing a dependency', which could be formalized in various ways.

A fact that has been the focus of much interest in the sentence processing literature is that in processing wh-dependencies, comprehenders do not appear to wait to see for certain that a gap is present before constructing the dependency and interpreting the new structure. Instead, they seem to anticipate the position of the gap in advance of the bottom-up information that would confirm it. Some evidence of this comes in the form of the 'filled-gap effect': when comprehenders encounter an object where they had predicted a gap, they demonstrate reaction time slow-downs (36) relative to a similar sentence without a filler (37) (Crain & Fodor, 1985; Stowe, 1986).

(36) My brother wanted to know who Ruth will bring <u>us</u> home to __ at Christmas.
(37) My brother wanted to know if Ruth will bring <u>us</u> home to Mom at Christmas.

The filled-gap effect shows that comprehenders do not wait for bottom-up evidence of a gap before constructing the dependency chain. Other evidence for this comes from studies that show processing disruption in advance of the gap when the verb is not plausible in combination with the filler as an object (38) (Traxler & Pickering, 1996; Garnsey, Tanenhaus, & Chapman, 1989). This shows that the construction of the dependency is not gated on the plausibility of the verb + object combination. More recently, small filled-gap effects have been observed in the subject position (Lee, 2004), suggesting that the parser attempts to fill the gap in the subject position first.

(38) That's the pistol/garage with which the heartless killer shot ___ the hapless man yesterday afternoon.

These findings suggest that encountering a filler triggers an expectation of the approximate type 'I expect to be interpreted in the first argument position I encounter[12].' Like the backwards anaphora case, this expectation cannot be obviously translated into an immediate prediction for a particular syntactic position.

Syntactic constraints on dependency formation govern filler gap constructions; although the filler and gap can be arbitrarily far away from each other, there are many domains that block the dependency from being well-formed, which are known as islands (Ross, 1967). Examples of island domains are subjects 0 and relative clauses (40).

(39) *What did the fact that Joan remembered __ surprise her grandchildren?
(40) *What did the agency fire the official that recommended __?

A number of studies have examined the online processing of wh-dependencies in island constraint contexts to determine whether comprehenders will initially erroneously build dependencies into islands, only to realize later that they are ill-formed. Overwhelmingly, the results have answered in the negative: comprehenders do not seem to construct dependencies that cross into island domains. For example, there is no filled-gap effect slow-down within a subject island (41) (Stowe, 1986) and there is also no filler-gap implausibility disruption in a relative clause island (42) (Traxler & Pickering, 1996). Although a few studies do find evidence for violating island constraints in

---

[12] Note that there exists a literature dedicated to determining exactly what are the dominant factors governing which positions are attempted to be filled, in particular what counts as the first relevant position for a gap (e.g. Aoshima et al. 2004); for me, it is just important that the expectation does not refer to a particular position in the structure.

dependency construction (Freedman & Forster, 1985; Clifton & Frazier, 1989; Pickering, Barton, & Shillcock, 1994), the authors argue that these findings can be accounted for by alternative interpretations (see also Phillips, 2006).

(41) The teacher asked what the silly story about <u>Greg's</u> older brother was supposed to mean.
(42) We like the book/city that the author who <u>wrote</u> unceasingly and with great dedication saw while waiting for a contract.

These findings lend more support to the hypothesis that it is the predictive nature of both backwards anaphora dependencies and wh-dependencies that makes processing so accurate with respect to the syntactic constraints governing them. Again, it seems plausible that the expectation engendered by the first element of both of these dependencies makes it possible to avoid the necessity of doing an error-prone retrieval operation when the second element is encountered (in this case the gap, which can be recognized in the position following the verb), as in the agreement attraction cases. However, the same question arises in the wh-dependency case of how exactly the predictive aspect of the process prevents erroneous dependencies to be formed between the filler and inaccessible gap positions. Similar to our discussion of backwards anaphora, several possible explanations are available. It could be the case that specific structural predictions for the gap are constantly being made and revised, which allow the parser to ignore intervening material that does not force revision of the predicted position (see Wagers & Phillips, 2009, for evidence that specific gap positions may be predicted well in advance). It could also be the case that the expectation is more abstract and does not take the form of making structural commitments predictively, but that searching for a gap position prospectively confers benefits not available in retrieval, as suggested above.

6.4 Conclusion

In this chapter, I have discussed behavioral findings in the processing of syntactic dependencies that suggest that syntactic prediction may make comprehension more accurate because it allows the parser to avoid the need for subsequent memory retrieval, which is prone to errors of interference. I first discussed evidence on the comprehension of subject-verb agreement demonstrating a selective pattern of interference from intervening material: interfering noun number made subject-verb number mismatches sound better, but did not make subject-verb number matches sound worse. I showed that this pattern could be accounted for by assuming that the number of the verb is predicted when the subject head was encountered, making the intervening material irrelevant for processing except when the prediction is violated. Violation of prediction is hypothesized to initiate a retrieval process that uses verb number as one cue to search for the licensing subject, a process prone to similarity-based interference errors on intervening material. I show that this hypothesis may be extended to account for why the processing of other dependencies that allow syntactic expectations seem not to be susceptible to interference from illicit elements (backwards anaphora, filler-gap dependencies) while dependencies that do not allow expectations are susceptible to interference (forwards pronominal anaphora). However, because the syntactic expectation allowed by backwards anaphora and filler-gap dependencies is not as limited as the agreement case, the mechanism by which the expectation allows irrelevant intervening candidates to be ignored is unclear. Therefore, our conjecture that predictive mechanisms make processing accurate for backwards anaphora and filler-gap dependencies remains somewhat speculative, and awaits confirmation from future work exploring the candidate accounts raised above.

7 Conclusion

7.1 Overview

In the previous chapters I have described evidence in support of the claim that predictive mechanisms play a role in particular aspects of language comprehension. In what follows, I discuss predictive mechanisms more generally—what counts as definitive evidence that contextual effects are due to predictive mechanisms, and how might pre-activation be implemented computationally—before turning to general conclusions.

7.2 Prediction or merely top-down effects?

As I discussed in the introduction, predictive mechanisms are those that involve activating or constructing representations on the basis of the context and without the benefit of external input. While the experiments here provide good evidence that context impacts early processing levels, none of them definitively show whether this impact takes place before the stimulus is presented. Predictive and top-down processes are typically examined by comparing the response to a critical stimulus in contexts that vary in the strength of their prediction for that stimulus. This is the approach that is taken by the experiments presented in this dissertation. However, interpretation of these data is continually plagued by the ambiguity of whether differences in the response really reflected prediction—a predictive change in the state of the system before the critical stimulus was presented—or whether the context only began to have its effect after the bottom-up information from the input began to make its way through the system.

The possibility of this alternative kind of account means that many findings that have been taken as evidence for 'prediction' may not actually unambiguously implicate pre-activation. For example, a very influential ERP study by Federmeier and Kutas (1999a) showed a reduced N400 to incongruous endings that shared semantic features with the predicted target ending (43).

(43) They wanted to make the hotel look like a tropical resort…
    a. …so along the driveway they planted rows of <u>palms</u>.
    b. …so along the driveway they planted rows of <u>pines</u>.
    c. …so along the driveway they planted rows of <u>roses</u>.

Although *pines* and *roses* are equally unexpected as endings, *pines* showed a smaller N400. Federmeier and Kutas argued that this was due to pre-activation of the semantic feature for 'tree' associated with *palms*, which then also served to prime *pines*. However, on the alternative accounts, this effect could equally be due to top-down prediction of semantic features based on the broader context only after the initial bottom-up information was processed, where again the contextual support for the 'tree' feature can act as a filter on the candidate representations activated by the input.

In Chapter 5, I pointed out that in order to use something like the earliness of the ELAN to make an argument for prediction, we would need better evidence about the timing of access of lexical information. A more definitive approach for demonstrating that representations are predictively activated or constructed prior to the supporting sensory input is for the experimental measure to also be gathered prior to the supporting sensory input. This kind of design has its own set of problems, but has recently been attempted with some success in both language and in other domains. In the following I discuss some of these approaches and their results.

Localization of pre-stimulus activity

One approach is to try to show pre-activation directly by identifying a candidate region where the target representations are stored and showing that this region is activated by other types of stimuli when they predict the target. For example, Gonzalo and Büchel (2004) used fMRI in humans to show that a tone that has been frequently paired with a face begins to activate the fusiform face area even in isolation, and Schlack and Albright (2007) similarly showed with single-cell recordings in monkeys that a static arrow that has been frequently paired with a particular direction of motion begins to activate MT, the visual area specialized in motion processing.

The difficulty of implementing this strategy for investigating prediction in language is that, given that we have no technique that we can use in humans with both very good precision in both time and space, we really need cases where processing the predictive context itself activates a very different area from processing of the stimulus it predicts. That's because we are looking for pre-activation in the same time-window that the context may still be being processed. The reason that the Gonzalo and Büchel (2004) study worked so well was that the context was in both a different modality and of a different sort (tones) than the thing being predicted (a face); since tones presumably activate the fusiform face area very little on their own, it was easy to pull out the increased activity there due to the prediction. However, if we consider the sentence context *I like my coffee with cream and* ___ which strongly predicts *sugar*, we can see that there is no sense in which the prediction for *sugar* should activate a completely different region from that activated by the preceding words.

For this approach to work for language, you would first have to have very discrete cortical organization of linguistic or conceptual representations, so that a prediction of one representation would have a clearly defined spatial locus different from another. For example, some investigators have suggested that nouns and verbs have different areas of posterior temporal cortex devoted to them (Bedny et al. 2006, 2008), so one could see whether contextual prediction of a noun activated the noun area more than the verb area prior to stimulus presentation. However, this case would be difficult to test naturalistically, as usually prediction of a noun vs. a verb comes following the presentation of other nouns and verbs in the sentence context, which would run up against the second problem of confusing the response to the context with response to the critical stimulus. Another possibility would be to make use of the proposed cortical sub-areas for animals and tools (Damasio et al., 1996); one could attempt to design sentence contexts that would predict one or the other (*I went out in the morning to walk the* __), but it would likely be difficult to do this for a whole materials set. Furthermore, for naturalistic sentence processing the context cannot be separated from the critical stimulus by a long gap, and thus the timescale of the BOLD response would make it difficult to distinguish the difference in timing of activation of the predicted area; on the other hand, it is unclear whether MEG could resolve the different spatial activation associated with animals vs tools. Therefore, although this is a very interesting approach to proving prediction, it is not obvious that it can be applied to language processing at this time.

Incongruency responses before the incongruency

A second approach is to show that the response to stimuli unlikely to have been predicted themselves is contingent on the predicted target. For example, Van Berkum and colleagues (Van Berkum et al., 2005) contrasted predicted and unpredicted targets of the form in 0.

(44) The burglar had no trouble finding the secret family safe. Of course it was behind a big painting (predicted) / bookcase (unpredicted).

What is interesting about this case is not the response to the target, but the response to the pre-target adjective. Dutch has gender-marking on nouns and adjectives, and since the gender was designed to differ between predicted and unpredicted ending (e.g., *painting* → neuter*, bookcase* → common gender), the realization of the preceding adjective also differed in gender-marking. Van Berkum and colleagues (2005) show an early positivity in the unpredicted condition following the gender-marking suffix of the adjective, prior to the onset of the noun. An account in which the noun is not activated until the bottom-up input for the noun is encountered would have to hold that this effect was due to contextual fit between the previous sentence and the gender-marked form of the adjective itself—for example, that in the past similar kinds of contexts are more frequently followed by a neuter-marked adjective than a common-marked adjective. However, given that the adjective is an optional modifier, and that in this study the adjective had to be fairly semantically vague in order to fit with both endings, it seems unlikely that the context would strongly predict particular adjectives, let alone their realizations. If not, the only remaining account for why the response would differ based on the gender-marking of the adjective is that it is inconsistent with a target noun that has been explicitly predicted. Wicha and colleagues (2003, 2004) and Otten and colleagues (2007) similarly report ERP effects of predicted gender features prior to the item predicted, although the timing and direction of these effects appears to be quite variable.

These results suggest at least that the formal gender features associated with the target lexical item can be pre-activated. Preliminary evidence from Szewczyk (2006) suggests that formal semantic features such as animacy can also be preactivated. Szewczyk shows an increased negativity on a preceding adjective with an agreement marker for animacy that mismatches the noun that is predicted, as in (45).

(45)
a. Together with her the children prepared <u>interesting-*inanimate*</u> spectacle. (congr.)
b. Together with her the children prepared <u>interesting-*animate*</u> professor. (incongr.)

Work by DeLong et al. (DeLong, Urbach, & Kutas, 2005) suggests that at least some of the phonological information associated with the target noun can also be pre-activated. They show an ERP effect in English on an article (*a* vs. *an*) phonologically contingent on the subsequent target noun (46).

(46)
a. (expected) The day was breezy so the boy went out to fly a <u>kite.</u>
b. (unexpected) The day was breezy so the boy went out to fly an <u>airplane</u>.

DeLong and colleagues find an increased negativity around 350-450 ms to the article (*a/an*) in the unexpected case, and the amplitude of the negativity is inversely correlated with the cloze probability. This difference may or may not reflect a typical N400 effect—it is within the same time-window, but seems to be more short-lived than most N400 effects—but critically, the fact that there is a difference at all between these conditions can only be plausibly explained if the phonological form of the subsequent noun has already been pre-activated.

Effects of context predictiveness prior to the stimulus

A third approach is to manipulate not the match between contextual prediction and stimulus identity, but rather the degree to which the context predicts a particular stimulus at all. If one context sets up a strong prediction for a particular ending, and another one does not, any differences between them in the pre-stimulus period could be taken to reflect mechanisms of prediction. In contrast to the first approach, which could only show effects of pre-activation, this approach could also show effects of pre-construction or pre-updating—any kind of predictive mechanism. One difficulty faced by this approach is in trying to match different contexts on every parameter except predictiveness. For example, sentence contexts that do not strongly constrain their ending tend to be less colorful and more vague than those that do. Therefore, differences between them might be taken to indicate differences in attention or similar higher-level processes, rather than predictive mechanisms.

Although a number of studies have manipulated the strength of contextual prediction, most have not examined the pre-stimulus period. One recent exception is work in progress by Suzanne Dikker and colleagues (p.c.) using MEG. Dikker used a novel design in which visual scene stimuli set up expectations for subsequent sentential stimuli. For example, a visual scene would either specify the kind of animal that Bill owns (a sheep) or leave it unknown. The subsequent sentence would read *Nick liked Mary's cow and also Bill's sheep*. Activity at *Bill's* could then examined for pre-stimulus prediction.

One difficulty inherent in this approach is pinpointing the exact moment at which a prediction is instantiated. Electrophysiological and neuroimaging techniques rely on averaging across large numbers of trials to beat the large signal to noise ratio in brain data, and this carries with it an assumption that the response of interest will be locked to a particular time-point across trials. However, in most naturalistic sentence stimuli, predictions unfold gradually across the sentence as more information is accrued. Even if the materials are similar enough that this timing would be relatively constant across trials, it is unclear where in time this would be. In the example sentence presented above, the prediction for *sheep* could have begun as early as at *and*, depending on the visual scene and the other materials in the experiment. Although Dikker and colleagues included other conditions that likely helped to constrain the timing of predictions, this example illustrates some of the challenges facing this approach.

7.3 Computational approaches to implementing pre-activation

It is pretty clear that predictive construction of upcoming structure should make use of the same system that is normally used for constructing that kind of structure, and facilitation in this case is definitionally a result of not having to do the work of constructing that structure later. However, what I have been calling predictive pre-activation—facilitation due to an expectation for a particular stored representation—could be realized in a number of ways. In this thesis I have remained agnostic about the implementation, but it is useful in looking forward to future research to consider several computational schemes that have been proposed for implementing effects of predictive pre-activation.

In general, effects of predictive context are observed as a reduction in activity when the stimulus is presented relative to when the stimulus is presented out of context, whether this activity is measured through single-electrode recordings in animals, ERPs, or fMRI. One class of models suggests that this is due to increased efficiency in processing in the neurons that represent the stimulus. This increased efficiency, in turn, can be realized in a number of different ways, as Grill-Spector, Henson, and Martin (2006) describe in an excellent review of the related phenomenon of repetition suppression. Since predictive pre-activation can be seen as a form of repetition (the representation is first activated by the prediction and then reactivated by the bottom-up information), the models can be straightforwardly extended to the predictive case. Figure 38 from Grill-Spector et al. (2006) illustrates three possible mechanisms by which the reduction in activity observed in predictive contexts might be realized.

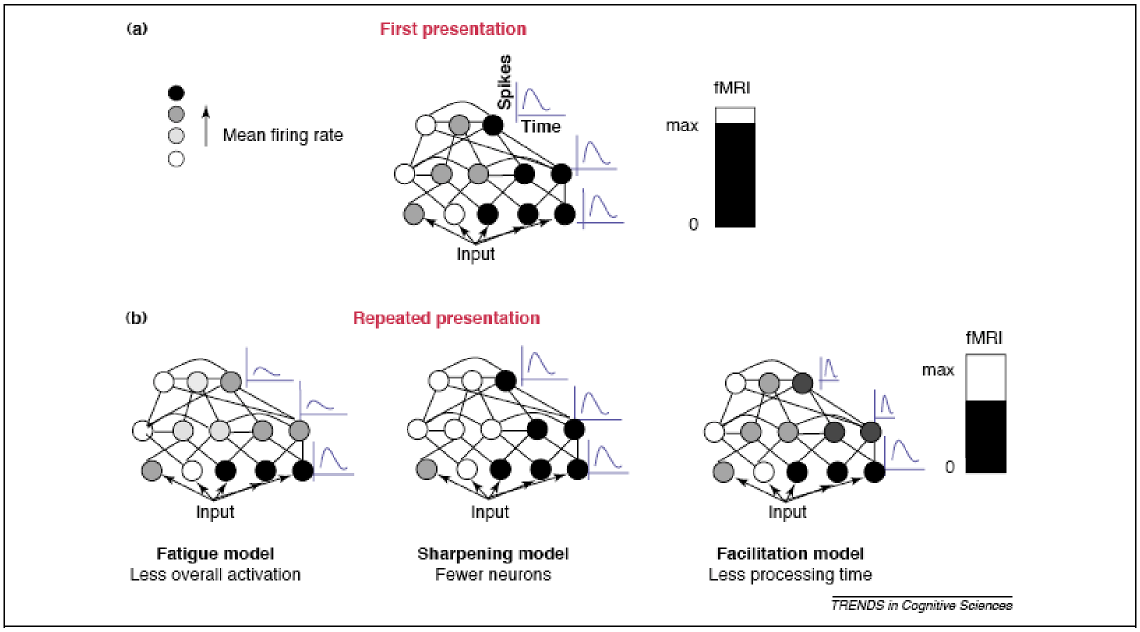


Figure 38. Models for repetition suppression from Grill-Spector et al. (2006). (a) illustrates that the visual stimulus is assumed to cause activity in the input layer (corresponding to early visual cortex) before being processed in a hierarchical sequence of stages on the initial presentation. The blue graphs indicate spiking (as a function of time) of the neurons with highest response at each stage (indicated by black circles). (b) illustrates various models for how repetition might lead to less BOLD activity.

First, when a stimulus is presented for a second time or after it has already been predictively pre-activated, all of the neurons that normally fire in response to the stimulus may show a reduced response, as if they were fatigued; hence this is labeled in the figure as the ‘Fatigue model’. Huber and O’Reilly (2003) have suggested that such a mechanism might be adaptive because it could prevent confusability of the responses to stimuli presented in rapid succession. It is not immediately obvious how the fatigue model can account for the behavioral facilitation in speed and accuracy observed for repeated or predicted stimuli, but Grill-Spector et al. (2006) note that it is possible that a mean reduction in activity could lead to greater neural synchrony, which might lead to faster transmission of the signal.

Second, when a stimulus is repeated or predicted, it could be that the neural response becomes more precise, such that the subset of neurons that code features that are irrelevant to identifying the stimulus show a reduction in activity on the second presentation, while the neurons optimally tuned to the stimulus respond with the same rate of activity (Desimone, 1996; Wiggs & Martin, 1998). Grill-Spector et al. refer to this as the ‘Sharpening model’. This model is appealing as it suggests that the representation itself becomes more efficient (relies on less neural activity to convey the needed information) and that this is what leads to more efficient behavior (although again it is unclear how this mechanism would lead the behavioral responses to be faster). The reduction in activity observed in fMRI and ERP is explained because a smaller population of neurons is firing, even though those that remain still fire at the same rate.

A third possibility is that the same neurons always respond at the same initial level of activity when a stimulus is presented, but that they fire for a shorter duration when the stimulus is repeated or predicted; this is referred to as the ‘Facilitation model’. The reduction in fMRI activity can be explained because fMRI integrates over several seconds of neural activity, so a response of 500 ms duration would result in a reduced signal compared to one of 1500 ms duration. Since the ERP can measure activity levels from ms to ms, reductions in the strength of the ERP response would be accounted for if the time window in which the reductions are observed is only the latter part of the time window of the stimulus response. This facilitation could be due to some form of synaptic potentiation between neurons that have recently responded together, resulting in information flowing through the network faster and processing being considered ‘completed’ at each stage faster.

As all three of these possibilities have been discussed so far, reduction of activity in response to a predicted stimulus may result as neurons become more fatigued, more selective, or finish sending their information faster because the stimulus representation was activated before the stimulus was presented. A very different perspective on this reduction of activity is that stimulus activity at a given level is actively suppressed as the activity can be ‘explained away’ by activated representations at higher levels.

A number of models assume that the system avoids redundancy of representation by always clearing or suppressing activity at a lower level of representation if this activity can be explained by the representation currently favored at the higher level (Rao & Ballard, 1999; Murray, Schrater, & Kersten, 2004; Friston, 2005). If we take the visual word case as an example, when there is no context, processing will proceed with simple visual units activating letter units, then letter units activating words. Whichever word is most highly activated will then suppress the letter units associated with that word, essentially clearing the activity at this level in preparation for the next word.

If the context predicts a word, the higher-level lexical representation is activated even before the stimulus, and therefore acts to suppress activity of units at lower levels that would be activated by presentation of that stimulus. Therefore, if the prediction is correct, almost no activity should result from the bottom-up input, as all of the units that would have reflected that activity will be selectively inhibited. This scheme could thus account for the reduced activity observed for stimuli presented in predictive contexts. The similarity between predictive effects and repetition priming could be accounted for if repetition is actually a special case of a predictive context; recent fMRI results suggesting that repetition priming is under strategic control support this (Summerfield et al., 2008).

One of the most important consequences of the 'explaining away' view is that it provides a means for the system to find the right solution if the prediction is only partially correct. To illustrate this with a toy example, if the context predicted the word 'CAT' to come up next, the units representing the letters 'C' 'A' and 'T' would be suppressed. If 'CAT' were then presented, there would be very little activity at the letter level, because the only units with perceptual support are suppressed. If 'MAT' was presented, only activity from the 'M' unit would make it up, telling the system that only the first letter was wrong (I am obviously glossing over many issues of letter order, etc.). In other words, whatever activity is not explained away can be used by the system to target the error and correct it. Rao and Ballard (1999) argue that this kind of mechanism can account for effects of 'end-stopping' observed in cells in primary visual cortex that fire more when one end of a line is presented in the cell's receptive field than when a line crosses through completely; the idea is that if a line extends on either side of the cell's receptive field, the representation of a line by other cells 'explains' the input to that cell, resulting in reduced activity. Murray and colleagues argue for such a mechanism on the basis of their fMRI finding that activity in primary visual cortex is reduced for visual input that has a higher-level interpretation (i.e., shapes vs. groups of lines matched on low-level visual parameters); this result can be accounted for if, when higher-level representations such as shapes can account for the input, lower-level units are suppressed (Murray, Kersten, Olshausen, Schrater, & Woods, 2002).

This view is strongly related to a longstanding literature in motor control that has emphasized the role of an efferent copy—the predicted consequence of a motor plan—in allowing the system to rapidly adjust to sensory feedback (e.g., Wolpert & Gharamani, 2000; Jeannerod, 2006). Rather than constructing the optimal motor plan in advance, which is difficult in changing contexts given the number of free parameters in movement, the system starts with an initial motor plan and generates an efferent copy that represents the expected sensory feedback from the next time step. The actual sensory feedback is then compared to the efferent copy, and the residual error between the predicted and actual feedback is used to correct the motor plan—an ‘analysis-by-synthesis’. Not only does this kind of ‘forward model’ make motor planning more accurate, but it provides a means for discriminating self-generated events (which will be part of the efferent copy) from external events (which will not). For example, these predictive mechanisms have been hypothesized to explain why a strong tickling sensation must usually be externally generated—why you can’t tickle yourself (Blakemore et al., 1999).

One way that the comparison step (between predicted and actual feedback) could be implemented is through suppression of the units representing the predicted input, and a number of auditory studies have demonstrated evidence of reductions in the neural response when the input is self-generated. In bats, activity in the auditory region of the lateral lemniscus in the midbrain is inhibited by vocalization (Suga & Schlegel, 1972), and in monkeys brain structures that are active during vocalization inhibit parts of auditory cortex (Müller-Preuss & Ploog, 1981). In humans, an early PET study showed that speaking modulated activity in auditory cortex (Paulesu, Frith, & Frackowiak 1993), and direct recordings from the human temporal lobe (Creutzfeldt, Ojemann, & Lettich, 1989) and MEG recordings (Numminen, Salmelin, & Hari, 1999) have also showed evidence of reduced auditory responses during speech. These results by themselves do not constitute strong evidence for a detailed forward model, as they could indicate a non-specific, across-the-board dampening of the auditory system that prevents auditory overload when the source is too close to the receptors. However, an MEG study by Martikainen and colleagues shows that an analogous reduction is also observed to tones when they are generated by a button-press from the participant compared to when they are externally generated (Martikainen et al., 2005). Furthermore, an MEG study by Houde and colleagues (Houde, Nagarajan, Sekihara, and Merzenich, 2002) provides evidence that the auditory suppression during speech is at least somewhat stimulus-specific: while the early auditory response to tones was similar whether the participant was actually speaking or just listening to a playback of their speech, the early auditory response to a vowel was much more strongly diminished by actual speech. Houde and colleagues also show in this and subsequent studies that less auditory suppression is observed during speech if the auditory feedback of the speech is altered, as would be expected if the amount of activity observed represents the error between the predicted and actual auditory feedback (Heinks-Maldonado, Mathalon, Gray, & Ford, 2005;

Heinks-Maldonado, Nagarajan, & Houde, 2006), and a recent fMRI study shows less activity for speaking with normal auditory feedback relative to altered feedback in posterior superior temporal cortex, a region implicated in auditory processing (Tourville, Reilly, & Guenther, 2008). Finally, several studies of audio-visual stimuli suggest that the auditory response is reduced when the visual stimuli predicts the timing and identity of the auditory stimulus (Oray, Lu, & Dawson, 2002; Klucharev et al., 2003; Van Wassenhove, Grant, & Poeppel, 2005; Stekelenburg & Vroomen, 2007).

The results of these studies are consistent with the hypothesis that predictions are instantiated through suppression of the predicted input and that this is the cause of the resulting reduction in activity often observed for predicted stimuli. At the same time, more evidence is needed to determine whether the suppressive mechanism implicated in motor control and specifically speech production is responsible for the predictive effects observed in normal language comprehension. Furthermore, even within the motor control literature there is some debate about whether it is actually the same units that represent the input at lower levels that are suppressed by the higher-level expectation, or whether suppression takes place over separate units specifically dedicated to representing the error between external feedback and internal predictions (e.g., Guenther, Ghosh, & Tourville, 2006). Some evidence for the latter scheme being implemented in one case comes from a recent fMRI study of audio-visual stimuli that showed increased activity in primary visual cortex for unpredicted audio-visual pairings even when it was the absence of a visual stimulus that was surprising (den Ouden, Friston, Daw, McIntosh, & Stephan, 2008); if the prediction only took the form of suppression of predicted units in primary visual cortex, increased activity when no stimulus is presented would be surprising.

7.4 General conclusions

In this dissertation I have joined a number of recent authors in arguing that contextual prediction is likely to play a central role in language comprehension, and I have presented experimental evidence that provides some initial constraints on the mechanisms that we might propose. I argue that pre-activation of stored lexical-semantic representations is at least one component of contextual effects on lexical processing, based on the MEG findings in Chapter 2 showing that contextual modulations of the neurophysiological response are similar in timing and distribution regardless of the degree of contextual fit or the amount of higher-level structure required by the context, as long as the context predicts the lexical or semantic content. The review of neuroimaging and neuropsychological data in Chapter 3 supports this claim, by showing that activity in a region believed to mediate storage of long-term lexical-semantic representations is affected by degree of contextual support. The earliness with which we show the ERP response to be affected by the syntactic requirements of the context may also provide constraints on models of predictive mechanisms, as I outline in Chapter 5.

I have also shown that, by including predictive and top-down mechanisms as part of 'normal' language processing rather than as optional add-ons to traditional feed-forward models, we gain insight into other elements of the system. By viewing the N400 paradigms as manipulating lexical predictability as well as contextual fit, and by assuming that lexical prediction should appear as a simple modulation of the same processes of access and integration necessary for language comprehension in general, I was able to make sense of variability in neural localization results that was otherwise puzzling. In the neural model I propose in Chapter 4, anterior inferior frontal cortex is linked not only to predictive pre-activation of lexical-semantic information in context, but is associated more generally with targeting information for retrieval from semantic memory in order to access critical relationships between representations or to meet the particular demands of a situation or task. A more posterior region of inferior frontal cortex that is linked to selecting between representations more generally may also be associated with selecting between candidate representations activated by the context through a top-down route vs. those activated by the input through a bottom-up route. The timing of the ELAN response discussed in Chapter 5 may ultimately contribute to constraints on models of the timecourse of feedforward processing in lexical access. If the generalization that I outline in Chapter 6 is correct, predictive mechanisms may also provide part of the solution to the question of how the parser can minimize the cost of memory interference for accurate processing of non-adjacent dependencies.

Finally, besides making specific contributions to our understanding of the timecourse and neuroanatomical correlates of language processing in context and of the potential advantages of predictive structure building for parsing, this work supports an approach to language processing which puts a strong emphasis on the impact of internally-generated hypotheses about the upcoming input in explaining how that input is interpreted. According to this approach, models of the timecourse of language processing will not be explanatory or predictive until they explicitly incorporate the timing with which prior knowledge about the constraints and likelihoods imposed by the context are implemented in addition to the timing with which the first feedforward volley of input information arrives at a given level of representation.

Bibliography

Addis, D. R., Wong, A. T., & Schacter, D. L. (2007). Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration. *Neuropsychologia*, *45*(7), 1363-1377.

Ainsworth-Darnell, K., Shulman, H. G., & Boland, J. E. (1998). Dissociating brain responses to syntactic and semantic anomalies: Evidence from event-related potentials. *Journal of Memory and Language*, *38*(1), 112-130.

Albright, T. D., & Stoner, G. R. (2002). Contextual influences on visual processing. *Annual Review of Neuroscience*, *25*, 339.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the Time Course of Spoken Word Recognition Using Eye Movements: Evidence for Continuous Mapping Models. *Journal of Memory and Language*, *38*, 419-439.

Almeida, D. (2009). *Form, Meaning and Context in Lexical Access: MEG and behavioral evidence.* Ph.D. dissertation, University of Maryland, College Park.

Aminoff, E., Schacter, D. L., & Bar, M. (2008). The cortical underpinnings of context-based memory distortion. *Journal of Cognitive Neuroscience*, *20*(12), 2226-2237.

Amunts, K., Schleicher, A., Burgel, U., Mohlberg, H., Uylings, H. B. M., & Zilles, K. (1999). Broca's region revisited: cytoarchitecture and intersubject variability. *The Journal of Comparative Neurology*, *412*(2).

Anderson, J. E., & Holcomb, P. J. (1995). Auditory and visual semantic priming using different stimulus onset asynchronies: an event-related brain potential study. *Psychophysiology*, *32*(2), 177-90.

Aoshima, S., Phillips, C., & Weinberg, A. (2004). Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language*, *51*(1), 23-54.

Austin, A., & Phillips, C. (2004). Rapid Syntactic Diagnosis: Separating Effects of Grammaticality and Expectancy. 17th Annual CUNY Conference on Human Sentence Processing, College Park, MD.

Badecker, W., & Lewis, R. (2007). A new theory and computational model of working memory in sentence production: agreement errors as failures of cue-based retrieval. 20th Annual CUNY Conference on Human Sentence Processing, La Jolla, CA.

Badecker, W., & Straub, K. (2002). The processing role of structural constraints on the interpretation of pronouns and anaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(4), 748.

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*, *45*(13), 2883-2901.

Badre, D., Poldrack, R. A., Pare-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron*, *47*(6), 907-18.

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, *11*(7), 280-289.

Bar, M. (2009). Predictions: a universal principle in the operation of the human brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1181.

Barrett, S. E., & Rugg, M. D. (1989). Event-related potentials and the semantic matching of faces. *Neuropsychologia*, *27*(7), 913-22.

Barrett, S. E., & Rugg, M. D. (1990). Event-related potentials and the semantic matching of pictures. *Brain and Cognition*, *14*(2), 201-12.

Baumgaertner, A., Weiller, C., & Büchel, C. (2002). Event-related fMRI reveals cortical sites involved in contextual sentence integration. *NeuroImage*, *16*, 736-745.

Bavelier, D., Corina, D., Jezzard, P., Padmanabhan, S., Clark, V. P., Karni, A., et al. (1997). Sentence Reading: a functional MRI study at 4 Tesla. *Journal of Cognitive Neuroscience*, *9*(5), 664-686.

Bedny, M., & Thompson-Schill, S. L. (2006). Neuroanatomically separable eVects of imageability and grammatical class during single-word comprehension. *Brain and Language*, *98*, 127-139.

Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A., & Saxe, R. (2008). Concepts Are More than Percepts: The Case of Action Verbs. *Journal of Neuroscience*, *28*(44), 11347.

Bedny, M., Hulbert, J. C., & Thompson-Schill, S. L. (2007). Understanding words in context: The role of Broca's area in word comprehension. *Brain research*, *1146*, 101-114.

Bedny, M., McGill, M., & Thompson-Schill, S. L. (2008). Semantic adaptation and competition during word comprehension. *Cerebral Cortex*, *18*(11), 2574.

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology*, *60*(4), 343-55.

Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., & Pernier, J. (1999). ERP Manifestations of Processing Printed Words at Different Psycholinguistic Levels: Time Course and Scalp Distribution. *Journal of Cognitive Neuroscience*, *11*(3), 235-260.

Bilenko, N. Y., Grindrod, C. M., Myers, E. B., & Blumstein, S. E. (2009). Neural Correlates of Semantic Competition during Processing of Ambiguous Words. *Journal of Cognitive Neuroscience*, 1-17.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. *Cerebral Cortex*.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Rao, S. M., & Cox, R. W. (1999). Conceptual processing during the conscious resting state: A functional MRI study. *Journal of Cognitive Neuroscience*, *11*(1), 80-93.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Springer, J. A., Kaufman, J. N., et al. (2000). Human Temporal Lobe Activation by Speech and Nonspeech Sounds. *Cerebral Cortex*, *10*(5), 512-528.

Binder, J. R., McKiernan, K. A., Parsons, M. E., Westbury, C. F., Possing, E. T., Kaufman, J. N., et al. (2003). Neural correlates of lexical access during visual word recognition. *Journal of Cognitive Neuroscience*, *15*(3), 372-93.

Binder, J. R., Medler, D. A., Desai, R., Conant, L. L., & Liebenthal, E. (2005). Some neurophysiological constraints on models of word naming. *Neuroimage*, *27*(3), 677-93.

Blakemore, S. J., Frith, C. D., & Wolpert, D. M. (1999). Spatio-temporal prediction modulates the perception of self-produced stimuli. *Journal of Cognitive Neuroscience*, *11*(5), 551-559.

Bloom, P. A., & Fischler, I. (1980). Completion norms for 329 sentence contexts. *Memory & Cognition*, *8*(6), 631-42.

Bock, J. K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, *23*(1), 45-93.

Bock, K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, *31*(1), 99-127.

Bock, K., & Eberhard, K. M. (1993). Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes*, *8*(1), 57-99.

Bokde, A. L. W., Tagamets, M. A., Friedman, R. B., & Horwitz, B. (2001). Functional interactions of the inferior frontal cortex during the processing of words and word-like stimuli. *Neuron*, *30*(2), 609-617.

Bolte, J., & Connine, C. M. (2004). Grammatical gender in spoken word recognition in German. *Perception and Psychophysics*, *66*(6), 1018-1032.

Brennan, J., & Pylkkänen, L. (2008). Processing events: Behavioral and neuromagnetic correlates of Aspectual Coercion. *Brain and Language*, *106*(2), 132-143.

Brennan, J., Nir, Y., Hasson, U., Malach, R., & Heeger, D. (submitted). Language in the looking glass: Linguistic processing during natural story listening.

Brown, C., & Hagoort, P. (1993). The Processing Nature of the N400: Evidence from Masked Priming. *Journal of Cognitive Neuroscience*, *5*(1), 34-44.

Buckner, R. L., Koutstaal, W., Schacter, D. L., & Rosen, B. R. (2000). Functional MRI evidence for a role of frontal and inferior temporal cortex in amodal components of priming. *Brain*, *123*(3), 620-640.

Buckner, R. L., Petersen, S. E., Ojemann, J. G., Miezin, F. M., Squire, L. R., & Raichle, M. E. (1995). Functional anatomical studies of explicit and implicit memory retrieval tasks. *Journal of Neuroscience*, *15*(1), 12-29.

Bullier, J., & Nowak, L. G. (1995). Parallel versus serial processing: new vistas on the distributed organization of the visual system. *Current Opinion in Neurobiology*, *5*(4), 497-503.

Camblin, C. C., Gordon, P. C., & Swaab, T. Y. (2007). The interplay of discourse congruence and lexical association during sentence processing: Evidence from ERPs and eye tracking. *Journal of Memory and Language*, *56*(1), 103-128.

Cappa, S. F., Perani, D., Schnur, T., Tettamanti, M., & Fazio, F. (1998). The effects of semantic category and knowledge type on lexical-semantic access: a PET study. *Neuroimage*, *8*(4), 350-9.

Cardillo, E. R., Aydelott, J., Matthews, P. M., & Devlin, J. T. (2004). Left inferior prefrontal cortex activity reflects inhibitory rather than facilitatory priming. *Journal of Cognitive Neuroscience*, *16*(9), 1552-61.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Clifton Jr, C., & Frazier, L. (1989). Comprehending sentences with long-distance dependencies. In Carlson, G. N. and Tanenhaus, M. K. *Linguistic structure in language processing* (pp. 273-317). Dordrecht: D. Reidel.

Clifton, C., Frazier, L., & Deevy, P. (1999). Feature manipulation in sentence comprehension. *Rivista di lingüística*, *11*(1), 11.

Clifton, C., Kennison, S. M., & Albrecht, J. E. (1997). Reading the Words Her, His, Him: Implications for Parsing Principles Based on Frequency and on Structure. *Journal of Memory and Language*, *36*(2), 276-292.

Connolly, J. F., & Phillips, N. A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal words of spoken sentences. *Journal of Cognitive Neuroscience*, *6*, 256-266.

Coulson, S., Federmeier, K. D., Van Petten, C., & Kutas, M. (2005). Right hemisphere sensitivity to word-and sentence-level context: Evidence from event-related brain potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 129–147.

Coulson, S., King, J. W., & Kutas, M. (1998). ERPs and Domain Specificity: Beating a Straw Horse. *Language and Cognitive Processes*, *13*(6), 653-672.

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*(3), 162-173.

Crain, S., & Fodor, J. D. (1985). How can grammars help parsers? In D. Dowty, L. Kartunnen, & A. Zwicky (Eds.), *Natural language parsing: Psychological, computational and theoretical perspectives* (pp. 94-127). Cambridge: Cambridge University Press.

Crescentini, C., Shallice, T., & Macaluso, E. (2009). Item Retrieval and Competition in Noun and Verb Generation: An fMRI Study. *Journal of Cognitive Neuroscience*, 1-18.

Creutzfeldt, O., Ojemann, G., & Lettich, E. (1989). Neuronal activity in the human lateral temporal lobe. II. Responses to the subjects own voice. *Experimental brain research,* *77*(3), 476.

Crinion, J. T., Lambon-Ralph, M. A., Warburton, E. A., Howard, D., & Wise, R. J. (2003). Temporal lobe regions engaged during normal speech comprehension. *Brain*, *126*(5), 1193-201.

Croxson, P. L., Johansen-Berg, H., Behrens, T. E. J., Robson, M. D., Pinsk, M. A., Gross, C. G., et al. (2005). Quantitative investigation of connections of the prefrontal cortex in the human and macaque using probabilistic diffusion tractography. *Journal of Neuroscience*, *25*(39), 8854-8866.

Curran, T., Tucker, D. M., Kutas, M., & Posner, M. I. (1993). Topography of the N400: brain electrical activity reflecting semantic expectancy. *Electroencephalography Clinical Neurophysiology*, *88*(3), 188-209.

Cutler, E. A., & Norris, D. (1979). Monitoring sentence comprehension. In W. Cooper & E. Walker (Eds.), *Sentence Processing: Psycholinguistic Studies presented to Merrill Garrett* . Hillsdale, NJ: Lawrence Earlbaum.

Damasio, A. R., & Damasio, H. (1994). Cortical systems for retrieval of concrete knowledge: The convergence zone framework. In C. Koch & J. L. Davis (Eds.), *Large-scale neuronal theories of the brain* (pp. 61-74). Cambridge, MA: MIT Press.

Damasio, H. (1991). Neuroanatomical correlates of the aphasias. In M. Sarno (Ed.), *Acquired aphasia* (Vol. 2, pp. 45–71). San Diego, CA: Academic.

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature*, *380*(6574), 499-505.

Dapretto, M., & Bookheimer, S. Y. (1999). Form and content: dissociating syntax and semantics in sentence comprehension. *Neuron*, *24*(2), 427-32.

Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *Journal of Neuroscience*, *23*(8), 3423–3431.

Davis, M. H., & Johnsrude, I. S. (2007). Hearing speech sounds: Top-down influences on the interface between audition and speech perception. *Hearing Research*, *229*(1-2), 132-147.

de Cheveigné, A., & Simon, J. Z. (2007). Denoising based on time-shift PCA. *Journal of Neuroscience Methods*, *165*(2), 297-305.

Deacon, D., Uhm, T. J., Ritter, W., Hewitt, S., & Dynowska, A. (1999). The lifetime of automatic semantic priming effects may exceed two seconds. *Cognitive Brain Research*, *7*(4), 465-472.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*(8), 1117-21.

Demb, J. B., Desmond, J. E., Wagner, A. D., Vaidya, C. J., Glover, G. H., & Gabrieli, J. D. (1995). Semantic encoding and retrieval in the left inferior prefrontal cortex: a functional MRI study of task difficulty and process specificity. *Journal of Neuroscience*, *15*(9), 5870-8.

den Ouden, H. E. M., Friston, K. J., Daw, N. D., McIntosh, A. R., & Stephan, K. E. (2008). A dual role for prediction error in associative learning. *Cerebral Cortex*.

Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences USA*, *93*, 13494-13499.

D'Esposito, M., Postle, B. R., Jonides, J., & Smith, E. E. (1999). The neural substrate and temporal dynamics of interference effects in working memory as revealed by event-related functional MRI. *Proceedings of the National Academy of Sciences USA, 96*, 7514-7519.

Devlin, J. T., Jamison, H. L., Matthews, P. M., & Gonnerman, L. M. (2004). Morphology and the internal structure of words. *Proceedings of the National Academy of Sciences USA*, *101*(41), 14984-8.

Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Wilson, J., Moss, H. E., et al. (2000). Susceptibility-Induced Loss of Signal: Comparing PET and fMRI on a Semantic Task. *Neuroimage*, *11*(6), 589-600.

Diaz, M. T., & Swaab, T. Y. (2007). Electrophysiological differentiation of phonological and semantic integration in word and sentence contexts. *Brain Research*, *1146*, 85-100.

Dien, J. (1998). Issues in the application of the average reference: Review, critiques, and recommendations. *Behavior research methods, instruments & computers*, *30*(1), 34-43.

Dikker, S., Rabagliati, H., & Pylkkänen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition*, *110*(3), 293-321.

Dobbins, I. G., & Wagner, A. D. (2005). Domain-general and domain-sensitive prefrontal mechanisms for recollecting events and detecting novelty. *Cerebral Cortex*, *15*(11), 1768-1778.

Dolan, R. J., & Fletcher, P. C. (1997). Dissociating prefrontal and hippocampal function in episodic memory encoding. *Nature, 388*(6642), 582-5.

Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, *92*(1-2), 145-77.

Duffy, S. A., Henderson, J. M., & Morris, R. K. (1989). Semantic facilitation of lexical access during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(5), 791-801.

Eberhard, K. (1997). The marked effect of number on subject-verb agreement. *Journal of Memory and Language*, *36*, 147-164.

Eberhard, K. M., Cutting, J. C., & Bock, K. (2005). Making syntax of sense: Number agreement in sentence production. *Psychological review, 112*(3), 531-558.

Elger, C. E., Grunwald, T., Lehnertz, K., Kutas, M., Helmstaedter, C., Brockhaus, A., et al. (1997). Human temporal lobe potentials in verbal learning and memory processes. *Neuropsychologia*, *35*(5), 657-67.

Elman, J. L., & McClelland, J. L. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language*, *27*(2), 143-165.

Farmer, T. A., Christiansen, M. H., & Monaghan, P. (2006). Phonological typicality influences on-line sentence comprehension. *Proceedings of the National Academy of Sciences USA*, *103*(32), 12203.

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

Federmeier, K. D., & Kutas, M. (1999a). A Rose by Any Other Name: Long-Term Memory Structure and Sentence Processing. *Journal of Memory and Language*, *41*, 469-495.

Federmeier, K. D., & Kutas, M. (1999b). Right words and left words: Electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research*, *8*, 373-392.

Federmeier, K. D., & Kutas, M. (2001). Meaning and Modality: Influences of Context, Semantic Memory Organization, and Perceptual Predictability on Picture Processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 202-224.

Federmeier, K. D., Van Petten, C., Schwartz, T. J., & Kutas, M. (2003). Sounds, words, sentences: Age-related changes across levels of language processing. *Psychology and aging*, *18*(4), 858-872.

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, *1146*, 75-84.

Fedorenko, E., & Kanwisher, N. (submitted). Neuroimaging of language: why hasn't a clearer picture emerged?

Ferstl, E. C., Rinck, M., & von Cramon, D. Y. (2005). Emotional and temporal aspects of situation model processing during text comprehension: an event-related fMRI study. *Journal of Cognitive Neuroscience*, *17*(5), 724-39.

Fischler, I., & Bloom, P. A. (1979). Automatic and Attentional Processes in the Effects of Sentence Contexts on World Recognition. *Journal of Verbal Learning and Verbal Behavior*, *18*(1), 1-20.

Fischler, I., Bloom, P. A., Childers, D. G., Roucos, S. E., & Perry, N. W. (1983). Brain Potentials Related to Stages of Sentence Verification. *Psychophysiology*, *20*(4), 400-409.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*(1), 116-124.

Franck, J., Vigliocco, G., & Nicol, J. (2002). Attraction in sentence production: The role of syntactic structure. *Language and Cognitive Processes, 17 (4)*, 371-404.

Franklin, M. S., Dien, J., Neely, J. H., Huber, E., & Waterson, L. D. (2007). Semantic priming modulates the N400, N300, and N400RP. *Clinical Neurophysiology*, *118*(5), 1053-1068.

Freedman, S. E., & Forster, K. I. (1985). The psychological status of overgenerated sentences. *Cognition*, *19*(2), 101-131.

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

Friederici, A. D., & Frisch, S. (2000). Verb Argument Structure Processing: The Role of Verb-Specific and Argument-Specific Information. *Journal of Memory and Language*, *43*(3), 476-507.

Friederici, A. D., & Kotz, S. A. (2003). The brain basis of syntactic processes: functional imaging and lesion studies. *Neuroimage*, *20*, 8-17.

Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: Early and late event-related brain potential effects. *Journal of Experimental Psychology-Learning Memory and Cognition*, *22*(5), 1219-1248.

Friederici, A. D., Hahne, A., & von Cramon, D. Y. (1998). First-pass versus second-pass parsing processes in a Wernicke's and a Broca's aphasic: electrophysiological evidence for a double dissociation. *Brain and Language*, *62*(3), 311-41.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, *1*(3), 183-92.

Friederici, A. D., Rüschemeyer, S. A., Hahne, A., & Fiebach, C. J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. *Cerebral Cortex*, *13*(2), 170-7.

Friederici, A. D., von Cramon, D. Y., & Kotz, S. A. (1999). Language related brain potentials in patients with cortical and subcortical left hemisphere lesions. *Brain*, *122 (6)*, 1033-47.

Frishkoff, G. A., Tucker, D. M., Davey, C., & Scherg, M. (2004). Frontal and posterior sources of event-related potentials in semantic comprehension. *Cognitive Brain Research*, *20*(3), 329-54.

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

Gabrieli, J. D. E., Desmond, J. E., Demb, J. B., Wagner, A. D., Stone, M. V., Vaidya, C. J., et al. (1996). Functional magnetic resonance imaging of semantic memory processes in the frontal lobes. *Psychological Science*, *7*(5), 278-283.

Gabrieli, J. D., Poldrack, R. A., & Desmond, J. E. (1998). The role of left prefrontal cortex in language and memory. *Proceedings of the National Academy of Sciences USA*,, *95*(3), 906-13.

Gagnepain, P., Chetelat, G., Landeau, B., Dayan, J., Eustache, F., & Lebreton, K. (2008). Spoken Word Memory Traces within the Human Auditory Cortex Revealed by Repetition Priming and Functional Magnetic Resonance Imaging. *Journal of Neuroscience*, *28*(20), 5281.

Ganis, G., Kutas, M., & Sereno, M. I. (1996). The Search for "Common Sense": An Electrophysiological Study of the Comprehension of Words and Pictures in Reading. *Journal of Cognitive Neuroscience*, *8*(2), 89-106.

Ganong 3rd, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, *6*(1), 110-25.

Gardiner, J. M., Craik, F. I., & Birtwistle, J. M. (1972). Retrieval cues and release from proactive inhibition. *Journal of Verbal Learning and Verbal Behavior*, *11*, 778-783.

Garnsey, S. M., Tanenhaus, M. K., & Chapman, R. M. (1989). Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research*, *18*(1), 51-60.

Giesbrecht, B., Camblin, C. C., & Swaab, T. Y. (2004). Separable effects of semantic priming and imageability on word processing in human cortex. *Cerebral Cortex*, *14*(5), 521-9.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91*(1), 1-67.

Giraud, A. L., Kell, C., Thierfelder, C., Sterzer, P., Russ, M. O., Preibisch, C., et al. (2004). Contributions of Sensory Input, Auditory Search and Verbal Comprehension to Cortical Activity during Speech Processing. *Cerebral Cortex*, *14*(3), 247-255.

Gitelman, D. R., Nobre, A. C., Sonty, S., Parrish, T. B., & Mesulam, M. M. (2005). Language network specializations: an analysis with parallel task designs and functional magnetic resonance imaging. *Neuroimage*, *26*(4), 975-85.

Gold, B. T., Balota, D. A., Jones, S. J., Powell, D. K., Smith, C. D., & Andersen, A. H. (2006). Dissociation of automatic and strategic lexical-semantics: functional magnetic resonance imaging evidence for differing roles of multiple frontotemporal regions. *Journal of Neuroscience*, *26*(24), 6523-32.

Gold, B. T., Balota, D. A., Kirchhoff, B. A., & Buckner, R. L. (2005). Common and Dissociable Activation Patterns Associated with Controlled Semantic and Phonological Processing: Evidence from fMRI Adaptation. *Cerebral Cortex*, *15*(9), 1438-1450.

Gonzalo, D., & Büchel, C. (2004). Audio-visual learning enhances responses to auditory stimuli in visual cortex. *Neuroimaging of visual cognition. Attention and performance XX.* Oxford: Oxford University Press.

Gorno-Tempini, M. L., Dronkers, N. F., Rankin, K. P., Ogar, J. M., Phengrasamy, L., Rosen, H. J., et al. (2004). Cognition and anatomy in three variants of primary progressive aphasia. *Annals of Neurology*, *55*(3), 335-346.

Gough, P. M., Nobre, A. C., & Devlin, J. T. (2005). Dissociating linguistic processes in the left inferior frontal cortex with transcranial magnetic stimulation. *Journal of Neuroscience*, *25*(35), 8010-6.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95-112.

Griffiths, T. D., Buchel, C., Frackowiak, R. S., & Patterson, R. D. (1998). Analysis of temporal structure in sound by the human brain. *Nature Neuroscience*, *1*(5), 422-427.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 14-23.

Grindrod, C. M., Bilenko, N. Y., Myers, E. B., & Blumstein, S. E. (2008). The role of the left inferior frontal gyrus in implicit semantic competition and selection: An event-related fMRI study. *Brain Research*, *1229*, 167-178.

Grossi, G. (2006). Relatedness proportion effects on masked associative priming: An ERP study. *Psychophysiology*, *43*(1), 21-30.

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language*, *96*(3), 280-301.

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

Hagoort, P. (2008). The fractionation of spoken language understanding by measuring electrical and magnetic brain signals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 1055-69.

Hagoort, P., Brown, C. M., & Swaab, T. Y. (1996). Lexical-semantic event-related potential effects in patients with left hemisphere lesions and aphasia, and patients with right hemisphere lesions without aphasia. *Brain*, *119 (Pt 2)*, 627-49.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, *304*(5669), 438-41.

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

Hahne, A., & Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research*, *13*(3), 339-356.

Halgren, E., Baudena, P., Heit, G., Clarke, J. M., Marinkovic, K., & Clarke, M. (1994). Spatio-temporal stages in face and word processing. I. Depth-recorded potentials in the human occipital, temporal and parietal lobes [corrected]. *Journal of Physiology: Paris*, *88*(1), 1-50.

Halgren, E., Baudena, P., Heit, G., Clarke, J. M., Marinkovic, K., Chauvel, P., et al. (1994). Spatio-temporal stages in face and word processing. 2. Depth-recorded potentials in the human frontal and Rolandic cortices. *Journal of Physiology: Paris*, *88*(1), 51-80.

Halgren, E., Dhond, R. P., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J. D., et al. (2002). N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. *Neuroimage*, *17*(3), 1101-16.

Hart Jr, J., & Gordon, B. (1990). Delineation of single-word semantic comprehension deficits in aphasia, with anatomical correlation. *Annals of Neurology*, *27*(3), 226-31.

Hartsuiker, R. J., Antón-Méndez, I., & van Zee, M. (2001). Object attraction in subject-verb agreement construction. *Journal of Memory and Language*, *45*(4), 546-572.

Haskell, T. R., & MacDonald, M. C. (2005). Constituent structure and linear order in language production: Evidence from subject verb agreement. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(5), 891.

Hassabis, D., & Maguire, E. A. (2009). The construction system of the brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1263.

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), 1383-400.

Häussler, J., & Bader, M. (2009). Agreement checking and number attraction in sentence comprehension: Insights from German relative clauses. *Travaux du Cercle Linguistique de Prague.*

Hawkins, J., & Blakeslee, S. (2004). *On Intelligence*. New York: Times Books.

Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., & Ford, J. M. (2005). Fine-tuning of auditory cortex during speech production. *Psychophysiology*, *42*(2), 180-190.

Heinks-Maldonado, T. H., Nagarajan, S. S., & Houde, J. F. (2006). Magnetoencephalographic evidence for a precise forward model in speech production. *Neuroreport*, *17*(13), 1375.

Helenius, P., Salmelin, R., Service, E., & Connolly, J. F. (1998). Distinct time courses of word and context comprehension in the left temporal cortex. *Brain*, *121(6)*, 1133-42.

Helenius, P., Salmelin, R., Service, E., & Connolly, J. F. (1999). Semantic cortical activation in dyslexic readers. *Journal of Cognitive Neuroscience*, *11*(5), 535-50.

Helenius, P., Salmelin, R., Service, E., Connolly, J. F., Leinonen, S., & Lyytinen, H. (2002). Cortical Activation during Spoken-Word Segmentation in Nonreading-Impaired and Dyslexic Adults. *Journal of Neuroscience*, *22*(7), 2936-2944.

Henson, R. N. A., Shallice, T., Josephs, O., & Dolan, R. J. (2002). Functional magnetic resonance imaging of proactive interference during spoken cued recall. *Neuroimage*, *17*(2), 543-558.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, *92*(1-2), 67-99.

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

Hill, H., Strube, M., Roesch-Ely, D., & Weisbrod, M. (2002). Automatic vs. controlled processes in semantic priming—differentiation by event-related potentials. *International Journal of Psychophysiology*, *44*(3), 197-218.

Hinojosa, J., Martín-Loeches, M., Casado, P., Muñoz, F., & Rubia, F. (2003). Similarities and differences between phrase structure and morphosyntactic violations in Spanish: An event-related potentials study. *Language and Cognitive Processes*, *18*(2), 113-142.

Hochstein, S., & Ahissar, M. (2002). View from the Top: Hierarchies and Reverse Hierarchies in the Visual System. *Neuron*, *36*(5), 791-804.

Holcomb, P. J. (1988). Automatic and attentional processing: an event-related brain potential analysis of semantic priming. *Brain and Language*, *35*(1), 66-85.

Holcomb, P. J., & McPherson, W. B. (1994). Event-related brain potentials reflect semantic priming in an object decision task. *Brain and Cognition*, *24*(2), 259-76.

Holcomb, P. J., & Neville, H. J. (1990). Auditory and Visual Semantic Priming in Lexical Decision: A Comparison Using Event-related Brain Potentials. *Language and Cognitive Processes*, *5*(4), 281-312.

Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: an MEG study. *Journal of Cognitive Neuroscience*, *14*(8), 1125-1138.

Huber, D. E., & O'Reilly, R. C. (2003). Persistence and accommodation in short-term priming and other perceptual paradigms: Temporal segregation through synaptic depression. *Cognitive Science*, *27*(3), 403-430.

Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Journal of Cognitive Neuroscience*, *18*(4), 665-79.

Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2007). Time course of semantic processes during sentence comprehension: An fMRI study. *Neuroimage*, *36*(3), 924-32.

Humphries, C., Love, T., Swinney, D. A., & Hickok, G. (2005). Response of Anterior Temporal Cortex to Prosodic and Syntactic Manipulations During Sentence Processing. *Human Brain Mapping*, *26*(2), 128-138.

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1-2), 101-44.

Iragui, V., Kutas, M., & Salmon, D. P. (1996). Event-related brain potentials during semantic categorization in normal aging and senile dementia of the Alzheimer's type. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *100*(5), 392-406.

Ischebeck, A., Indefrey, P., Usui, N., Nose, I., Hellwig, F., & Taira, M. (2004). Reading in a regular orthography: an FMRI study investigating the role of visual familiarity. *Journal of Cognitive Neuroscience*, *16*(5), 727-41.

Isel, F., Hahne, A., Maess, B., & Friederici, A. D. (2007). Neurodynamics of sentence interpretation: ERP evidence from French. *Biological psychology*, *74*(3), 337-346.

January, D., Trueswell, J. C., & Thompson-Schill, S. L. (2009). Co-localization of Stroop and Syntactic Ambiguity Resolution in Broca's Area: Implications for the Neural Basis of Sentence Processing. *Journal of Cognitive Neuroscience*.

Jeannerod, M. (2006). *Motor cognition: What actions tell the self.* Oxford University Press, USA.

Johnson, B. W., & Hamm, J. P. (2000). High-density mapping in an N400 paradigm: evidence for bilateral temporal lobe generators. *Clinical Neurophysiology*, *111*(3), 532-45.

Jonides, J., & Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience*, *139*(1), 181-193.

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2007). The mind and brain of short-term memory. *Annual Reviews of Psychology, 59,* 193-224.

Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A., & Reuter-Lorenz, P. A. (1998). Inhibition in verbal working memory revealed by brain activation. *Proceedings of the National Academy of Sciences USA, 95*, 8410-8413.

Kaan, E. (2002). Investigating the effects of distance and number interference in processing subject-verb dependencies: an ERP study. *Journal of Psycholinguistic Research*, *31*(2), 165-93.

Kaan, E., & Swaab, T. Y. (2003). Electrophysiological evidence for serial sentence processing: a comparison between non-preferred and ungrammatical continuations. *Cognitive Brain Research*, *17*(3), 621-35.

Kang, A. M., Constable, R. T., Gore, J. C., & Avrutin, S. (1999). An event-related fMRI study of implicit phrase-level syntactic and semantic processing. *Neuroimage, 10*(5), 555-61.

Kapur, S., Craik, F. I. M., Tulving, E., Wilson, A. A., Houle, S., & Brown, G. M. (1994). Neuroanatomical correlates of encoding in episodic memory: levels of processing effect. *Proceedings of the National Academy of Sciences USA*, *91*(6), 2008-2011.

Kazanina, N., Lau, E. F., Lieberman, M., Yoshida, M., & Phillips, C. (2007). The effect of syntactic constraints on the processing of backwards anaphora. *Journal of Memory and Language*, *56*(3), 384-409.

Kennison, S. M. (2003). Comprehending the pronouns her, him, and his: Implications for theories of referential processing. *Journal of Memory and Language*, *49*(3), 335-352.

Kertesz, A. (1979). *Aphasia and associated disorders: taxonomy, localization, and recovery*. New York: Grune & Stratton.

Kho, K. H., Indefrey, P., Hagoort, P., van Veelen, C. W. M., van Rijen, P. C., & Ramsey, N. F. (2007). Unimpaired sentence comprehension after anterior temporal cortex resection. *Neuropsychologia*.

Kiefer, M. (2002). The N400 is modulated by unconsciously perceived masked words: further evidence for an automatic spreading activation account of N400 priming effects. *Cognitive Brain Research*, *13*(1), 27-39.

Kiehl, K. A., Laurens, K. R., & Liddle, P. F. (2002). Reading anomalous sentences: an event-related fMRI study of semantic processing. *Neuroimage*, *17*(2), 842-50.

Kimball, J., & Aissen, J. (1971). I think, you think, he think. *Linguistic Inquiry*, 2, 241-246.

Klucharev, V., Möttönen, R., & Sams, M. (2003). Electrophysiological indicators of phonetic and non-phonetic multisensory interactions during audiovisual speech perception. *Cognitive Brain Research*, *18*(1), 65-75.

Kluender, R., & Kutas, M. (1993a). Subjacency as a processing phenomenon. *Language and Cognitive Processes*, *8*(4), 573-633.

Kluender, R., & Kutas, M. (1993b). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience*, *5*(2), 196-214.

Knill, D. C., & Richards, W. (1996). *Perception as Bayesian inference*. New York: Cambridge University Press.

Kojima, T., & Kaga, K. (2003). Auditory lexical-semantic processing impairments in aphasic patients reflected in event-related potentials (N400). *Auris Nasus Larynx*, *30*(4), 369-78.

Kotz, S. A., Cappa, S. F., von Cramon, D. Y., & Friederici, A. D. (2002). Modulation of the lexical-semantic network by auditory semantic priming: an event-related functional MRI study. *Neuroimage*, *17*(4), 1761-72.

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.

Kuhl, B., & Wagner, A. D. (2009). Forgetting and retrieval. In G. G. Berntson & J. T. Cacioppo (Eds.), *Handbook of Neurosciences for the Behavioral Sciences*. John Wiley and Sons.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, 23-49.

Kuperberg, G. R., Holcomb, P. J., Sitnikova, T., Greve, D., Dale, A. M., & Caplan, D. (2003). Distinct Patterns of Neural Modulation during the processing of Conceptual and Syntactic Anomalies. *Journal of Cognitive Neuroscience*, *15*(2), 272-293.

Kuperberg, G. R., Sitnikova, T., & Lakshmanan, B. M. (2008). Neuroanatomical distinctions within the semantic system during sentence comprehension: Evidence from functional magnetic resonance imaging. *Neuroimage*, *40*(1), 367-388.

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

Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition*, *11*(5), 539–550.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*(5947), 161-3.

Kutas, M., Van Petten, C., & Kluender, R. (2006). Psycholinguistics electrified II (1999–2005). In M. J. Traxler & M. A. Gernsbacher (Eds.), *The handbook of psycholinguistics* (2nd ed., pp. 659-724). San Diego, CA: Elsevier.

Ladefoged, P., & Broadbent, D. E. (1957). Information Conveyed by Vowels. *The Journal of the Acoustical Society of America*, *29*, 98-104.

Lahar, C. J., Tun, P. A., & Wingfield, A. (2004). Sentence-Final Word Completion Norms for Young, Middle-Aged, and Older Adults. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *59*(1), 7-10.

Lamme, V. A. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience*, *15*(2), 1605-1615.

Laszlo, S., & Federmeier, K. D. (2008). Minding the PS, queues, and PXQs: Uniformity of semantic processing across multiple stimulus types. *Psychophysiology*, *45*(3), 458-466.

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

Lau, E., Almeida, D., AbdulSabur, N., Braun, A., & Poeppel, D. (2009). Fractionating the N400 effect with simultaneous MEG and EEG. Poster presented at the 16th Cognitive Neuroscience Society Meeting, San Francisco, CA.

Lau, E., Rozanova, K., & Phillips, C. (2007). Syntactic Prediction and Lexical Surface Frequency Effects in Sentence Processing, *University of Maryland Working Papers in Linguistics*.

Lau, E., Stroud, C., Plesch, S., & Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and Language*, *98*(1), 74-88.

Lee, M. W. (2004). Another look at the role of empty categories in sentence processing (and grammar). *Journal of psycholinguistic research*, *33*(1), 51-73.

Lee, M. W., & Williams, J. (submitted). The role of grammatical constraints in intra-sentential pronoun resolution.

Lee, T. S., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *Journal of the Optical Society of America A*, *20*(7), 1434-1448.

Lee, T. S., & Nguyen, M. (2001). Dynamics of subjective contour formation in the early visual cortex. *Proceedings of the National Academy of Sciences*, *98*(4), 1907.

Lee, T. S., Mumford, D., Romero, R., & Lamme, V. A. F. (1998). The role of the primary visual cortex in higher level vision. *Vision Research*, *38*(15-16), 2429-2454.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science: A Multidisciplinary Journal*, *29*(3), 375-419.

Lindenberg, R., & Scheef, L. (2007). Supramodal language comprehension: Role of the left temporal lobe for listening and reading. *Neuropsychologia*, *45*(10), 2407-2415.

Llinás, R. R. (2001). *I of the vortex: from neurons to self*. Cambridge, MA: MIT Press.

Lobeck, A. (1995). *Ellipsis: Functional heads, licensing, and identification.* Oxford University Press, USA.

Luck, S. (2005). *An introduction to the event-related potential technique.* Cambridge, MA: MIT Press.

Lukatela, G., Kostic, A., Feldman, L., & Turvey, M. (1983). Grammatical priming of inflected nouns. *Memory & Cognition*, *11*(1), 59-63.

Lukatela, G., Moraca, J., Stojnov, D., Savic, M. D., Katz, L., & Turvey, M. T. (1982). Grammatical priming effects between pronouns and inflected verb forms. *Psychological Research*, *44*(4), 297-311.

Maess, B., Herrmann, C. S., Hahne, A., Nakamura, A., & Friederici, A. D. (2006). Localizing the distributed language network responsible for the N400 measured by MEG during auditory sentence processing. *Brain Research*, *1096*(1), 163-72.

Marí-Beffa, P., Valdés, B., Cullen, D. J. D., Catena, A., & Houghton, G. (2005). ERP analyses of task effects on semantic processing from words. *Cognitive Brain Research*, *23*(2-3), 293-305.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177-190.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1-2), 71-102.

Martikainen, M. H., Kaneko, K., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex*, *15*(3), 299-302.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, *58*, 25-45.

Mason, R. A., & Just, M. A. (2007). Lexical ambiguity in sentence comprehension. *Brain research*, *1146*, 115-127.

Mason, R. A., Just, M. A., Keller, T. A., & Carpenter, P. A. (2003). Ambiguity in the brain: What brain imaging reveals about the processing of syntactically ambiguous sentences. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *29*(6), 1319-1338.

Matsumoto, A., Iidaka, T., Haneda, K., Okada, T., & Sadato, N. (2005). Linking semantic priming effect in functional MRI and event-related potentials. *Neuroimage*, *24*(3), 624-34.

Mazoyer, B. M., Dehaene, S., Tzourio, N., Frak, V., Murayama, N., Cohen, L., et al. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience*, *4*(467-479).

McCarthy, G., Nobre, A. C., Bentin, S., & Spencer, D. D. (1995). Language-related field potentials in the anterior-medial temporal lobe: I. Intracranial distribution and neural generators. *Journal of Neuroscience*, *15*(2), 1080-9.

McClelland, J. L. & Elman, J. L. (1986). The TRACE Model of Speech Perception. *Cognitive Psychology,* 18, 1-86.

McDermott, K. B., Petersen, S. E., Watson, J. M., & Ojemann, J. G. (2003). A procedure for identifying regions preferentially activated by attention to semantic and phonological relations using functional magnetic resonance imaging. *Neuropsychologia*, *41*(3), 293-303.

McElree, B. (2006). Accessing recent events. In Ross, B. H. (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 46, p. 155). San Diego: Academic Press.

McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, *48*(1), 67-91.

Mechelli, A., Gorno-Tempini, M. L., & Price, C. J. (2003). Neuroimaging studies of word and pseudoword reading: consistencies, inconsistencies, and limitations. *Journal of Cognitive Neuroscience*, *15*(2), 260-71.

Mechelli, A., Josephs, O., Lambon Ralph, M. A., McClelland, J. L., & Price, C. J. (2007). Dissociating stimulus-driven semantic and phonological effect during reading and naming. *Human Brain Mapping*, *28*(3), 205-17.

Menenti, L., Petersson, K. M., Scheeringa, R., & Hagoort, P. (2009). When elephants fly: Differential sensitivity of right and left inferior frontal gyri to discourse and world knowledge. *Journal of Cognitive Neuroscience*, 1-11.

Metzler, C. (2001). Effects of left frontal lesions on the selection of context-appropriate meanings. *Neuropsychology*, *15*(3), 315.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Reviews of Neuroscience, 24,* 167-202.

Müller-Preuss, P., & Ploog, D. (1981). Inhibition of auditory cortical neurons during phonation. *Brain research*, *215*(1-2), 61.

Mummery, C. J., Patterson, K., Price, C. J., Ashburner, J., Frackowiak, R. S. J., & Hodges, J. R. (2000). A voxel-based morphometry study of semantic dementia: Relationship between temporal lobe atrophy and semantic memory. *Annals of Neurology*, *47*(1), 36-45.

Mummery, C. J., Patterson, K., Wise, R. J. S., Vandenbergh, R., Price, C. J., & Hodges, J. R. (1999). Disrupted temporal lobe connections in semantic dementia. *Brain*, *122*(1), 61-73.

Munte, T. F., Matzke, M., & Johannes, S. (1997). Brain activity associated with syntactic incongruencies in words and pseudo-words. *Journal of Cognitive Neuroscience*, *9*(3), 318-329.

Murray, S. O., Kersten, D., Olshausen, B. A., Schrater, P., & Woods, D. L. (2002). Shape perception reduces activity in human primary visual cortex. *Proceedings of the National Academy of Sciences*, *99*(23), 15164-15169.

Murray, S. O., Schrater, P., & Kersten, D. (2004). Perceptual grouping and the interactions between visual cortical areas. *Neural Networks*, *17*(5-6), 695-705.

Nakao, M., & Miyatani, M. (2007). Dissociation of semantic and expectancy effects on N400 using Neely's version of semantic priming paradigm: N400 reflects post-lexical integration. In T. Sakamoto (Ed.), *Communicating Skills of Intention* (pp. 21-31). Tokyo: Hituzi Syobo.

Nee, D. E., & Jonides, J. (2008). Neural correlates of access to short-term memory. *Proceedings of the National Academy of Sciences*, *105*(37), 14228.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology General*, *106*, 226-54.

Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: roles of prospective prime-generated expectancies and retrospective semantic matching. *J Exp Psychol Learn Mem Cogn*, *15*(6), 1003-19.

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.

Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, *3*(2), 151-165.

Newman, A. J., Pancheva, R., Ozawa, K., Neville, H. J., & Ullman, M. T. (2001). An event-related fMRI study of syntactic and semantic violations. *Journal of Psycholinguistic Research*, *30*(3), 339-64.

Nicol, J. L., Forster, K. I., & Veres, C. (1997). Subject–verb agreement processes in comprehension. *Journal of Memory and Language*, *36*(4), 569-587.

Nicol, J., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research*, *18*(1), 5-19.

Nobre, A. C., & McCarthy, G. (1994). Language-related ERPS - scalp distributions and modulation by word type and semantic priming. *Journal of Cognitive Neuroscience*, *6*(3), 233-255.

Nobre, A. C., & McCarthy, G. (1995). Language-related field potentials in the anterior-medial temporal lobe: II. Effects of word type and semantic priming. *Journal of Neuroscience*, *15*(2), 1090-8.

Nobre, A. C., Allison, T., & McCarthy, G. (1994). Word recognition in the human inferior temporal lobe. *Nature*, *372*(6503), 260-3.

Noppeney, U., & Price, C. J. (2004). An FMRI study of syntactic adaptation. *Journal of Cognitive Neuroscience*, *16*(4), 702-13.

Noppeney, U., Josephs, O., Hocking, J., Price, C. J., & Friston, K. J. (2008). The Effect of Prior Visual Information on Recognition of Speech and Sounds. *Cerebral Cortex*, *18*(3), 598.

Noppeney, U., Patterson, K., Tyler, L. K., Moss, H., Stamatakis, E. A., Bright, P., et al. (2007). Temporal lobe lesions and semantic impairment: a comparison of herpes simplex virus encephalitis and semantic dementia. *Brain*, *130*(4), 1138.

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: feedback is never necessary. *Behavioral and Brain Sciences*, *23*(3), 299-325.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(3), 263-281.

Numminen, J., Salmelin, R., & Hari, R. (1999). Subject's own speech reduces reactivity of the human auditory cortex. *Neuroscience letters*, *265*(2), 119-122.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97-113.

Olichney, J. M., Morris, S. K., et al. (2002). Abnormal verbal event related potentials in mild cognitive impairment and incipient Alzheimer's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, *73*(4), 377-384.

Olichney, J. M., Riggins, B. R., et al. (2002). Reduced Sensitivity of the N400 and Late Positive Component to Semantic Congruity and Word Repetition in Left Temporal Lobe Epilepsy. *Clinical Electroencephalography*, *33*(3), 111-118.

Olichney, J. M., Van Petten, C., Paller, K. A., Salmon, D. P., Iragui, V. J., & Kutas, M. (2000). Word repetition in amnesia. Electrophysiological measures of impaired and spared memory. *Brain*, *123*(9), 1948-63.

Oray, S., Lu, Z. L., & Dawson, M. E. (2002). Modification of sudden onset auditory ERP by involuntary attention to visual stimuli. *International Journal of Psychophysiology*, *43*(3), 213-224.

Orfanidou, E., Marslen-Wilson, W. D., & Davis, M. H. (2006). Neural response suppression predicts repetition priming of spoken words and pseudowords. *Journal of Cognitive Neuroscience*, *18*(8), 1237-1252.

Orgs, G., Lange, K., Dombrowski, J. H., & Heil, M. (2008). N400-effects to task-irrelevant environmental sounds: Further evidence for obligatory conceptual processing. *Neuroscience Letters*, *436*(2), 133-137.

Osterhout, L. (1997). On the brain response to syntactic anomalies: Manipulations of word position and word class reveal individual differences. *Brain and Language*, *59*(3), 494-522.

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

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, *34*, 739-773.

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

Otten, M., Nieuwland, M. S., & Van Berkum, J. J. A. (2007). Great expectations: Specific lexical anticipation influences the processing of spoken language. *BMC neuroscience*, *8*(1), 89.

Oztekin, I., Curtis, C. E., & McElree, B. (2009). The Medial Temporal Lobe and the Left Inferior Prefrontal Cortex Jointly Support Interference Resolution in Verbal Working Memory. *Journal of Cognitive Neuroscience*, 1-13.

Paradis, A. L., Cornilleau-Peres, V., Droulez, J., Van De Moortele, P. F., Lobel, E., Berthoz, A., et al. (2000). Visual perception of motion and 3-D structure from motion: an fMRI study. *Cerebral Cortex*, *10*(8), 772-783.

Paulesu, E., Frith, C. D., & Frackowiak, R. S. J. (1993). The neural correlates of the verbal component of working memory. *Nature, 362*(6418), 342-5.

Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, *41*(3), 427-456.

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.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature*, *331*(6157), 585-9.

Petrides, M., & Pandya, D. N. (2002). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *European Journal of Neuroscience*, *16*(2), 291.

Phillips, C. (2006). The real-time status of island phenomena. *Language*, *82*(4), 795-823.

Phillips, C., Wagers, M., & Lau, E. (submitted). Grammatical illusions and selective fallibility in real-time language comprehension.

Pickering, M. J., Barton, S., & Shillcock, R. (1994). Unbounded dependencies, island constraints and processing complexity. In C. Clifton Jr, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 199–224). London: Lawrence Earlbaum.

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage*, *10*(1), 15-35.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55-85). Hillsdale, NJ: Erlbaum.

Price, C. J., Wise, R. J. S., & Frackowiak, R. S. J. (1996). Demonstrating the implicit processing of visually presented words and pseudowords. *Cerebral Cortex*, *6*(1), 62-70.

Price, C. J., Wise, R. J. S., Watson, J. D. G., Patterson, K., Howard, D., & Frackowiak, R. S. J. (1994). Brain activity during reading: The effects of exposure duration and task. *Brain*, *117*(6), 1255.

Pugh, K. R., Shaywitz, B. A., Shaywitz, S. E., Constable, R. T., Skudlarski, P., Fulbright, R. K., et al. (1996). Cerebral organization of component processes in reading. *Brain*, *119(4)*, 1221-38.

Pylkkänen, L., & Marantz, A. (2003). Tracking the time course of word recognition with MEG. *Trends in Cognitive Sciences*, *7*(5), 187-189.

Pylkkänen, L., & McElree, B. (2007). An MEG Study of Silent Meaning. *Journal of Cognitive Neuroscience*, *19*(11), 1905-1921.

Pylkkänen, L., Martin, A. E., McElree, B., & Smart, A. (2009). The Anterior Midline Field: Coercion or decision making? *Brain and Language*, *108*(3), 184-190.

Pylkkänen, L., Stringfellow, A., & Marantz, A. (2002). Neuromagnetic evidence for the timing of lexical activation: An MEG component sensitive to phonotactic probability but not to neighborhood density. *Brain and Language*, *81*(1-3), 666-678.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.

Raichle, M. E., Fiez, J. A., Videen, T. O., MacLeod, A. M. K., Pardo, J. V., Fox, P. T., et al. (1994). Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex*, *4*(1), 8-26.

Randall, B., & Marslen-Wilson, W. D. (1998). The relationship between lexical and syntactic processing. In *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society.* Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*, 79-87.

Rao, R. P. N., Olshausen, B. A., & Lewicki, M. S. (2002). *Probabilistic models of the brain*. MIT Press.

Reddy, L., & Kanwisher, N. (2006). Coding of visual objects in the ventral stream. *Current Opinion in Neurobiology*, *16*(4), 408-414.

Rissman, J., Eliassen, J. C., & Blumstein, S. E. (2003). An event-related FMRI investigation of implicit semantic priming. *Journal of Cognitive Neuroscience*, *15*(8), 1160-75.

Robertson, L. C., & Lamb, M. R. (1991). Neuropsychological contributions to theories of part/whole organization. *Cognitive Psychology*, *23*(2), 299-330.

Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex*, *15*(8), 1261-9.

Rogalsky, C., & Hickok, G. (2009). Selective Attention to Semantic and Syntactic Features Modulates Sentence Processing Networks in Anterior Temporal Cortex. *Cerebral Cortex, 19*(4):786-96

Rogalsky, C., Saberi, K., & Hickok, G. (2009). Temporal and structural contributions to activation of anterior temporal sentence processing regions: an fMRI study. *15th Annual Meeting of the Organization for Human Brain Mapping, San Francisco, CA*.

Rosenfelt, L., Barkley, C., Kellogg, M. K., Kluender, R., & Kutas, M. (2009). No ERP Evidence for Automatic First-Pass Parsing. Paper presented at the 22nd Annual CUNY Conference on human sentence processing, Davis, CA.

Roskies, A. L., Fiez, J. A., Balota, D. A., Raichle, M. E., & Petersen, S. E. (2001). Task-dependent modulation of regions in the left inferior frontal cortex during semantic processing. *Journal of Cognitive Neuroscience*, *13*(6), 829-43.

Ross, J. R. (1967). Constraints on variables in syntax. Ph.D. Dissertation, Massachusetts Institute of Technology.

Rossell, S. L., Price, C. J., & Nobre, A. C. (2003). The anatomy and time course of semantic priming investigated by fMRI and ERPs. *Neuropsychologia*, *41*(5), 550-64.

Rossi, A. F., Desimone, R., & Ungerleider, L. G. (2001). Contextual modulation in primary visual cortex of macaques. *Journal of Neuroscience*, *21*(5), 1698.

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

Rugg, M. D. (1985). The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology*, *22*(6), 642-7.

Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high- and low-frequency words. *Memory & Cognition*, *18*(4), 367-79.

Rüschemeyer, S. A., Fiebach, C. J., Kempe, V., & Friederici, A. D. (2005). Processing lexical semantic and syntactic information in first and second language: fMRI evidence from German and Russian. *Human Brain Mapping*, *25*(2), 266-86.

Rüschemeyer, S. A., Zysset, S., & Friederici, A. D. (2006). Native and non-native reading of sentences: an fMRI experiment. *Neuroimage*, *31*(1), 354-65.

Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M. S., et al. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences USA*, *105*(46), 18035.

Schlack, A., & Albright, T. D. (2007). Remembering Visual Motion: Neural Correlates of Associative Plasticity and Motion Recall in Cortical Area MT. *Neuron*, *53*(6), 881-890.

Schuberth, R. E., & Eimas, P. D. (1977). Effects of Context on the Classification of Words and Nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(1), 27-36.

Scott, S. K. (2005). Auditory processing—speech, space and auditory objects. *Current Opinion in Neurobiology*, *15*(2), 197-201.

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, *123*(12), 2400-6.

Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: eye movements and event-related potentials. *Trends in Cognitive Sciences*, *7*(11), 489-493.

Sereno, S. C., Rayner, K., & Posner, M. I. (1998). Establishing a time-line of word recognition: evidence from eye movements and event-related potentials. *Neuroreport*, *9*(10), 2195-200.

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

Simos, P. G., Basile, L. F., & Papanicolaou, A. C. (1997). Source localization of the N400 response in a sentence-reading paradigm using evoked magnetic fields and magnetic resonance imaging. *Brain Research*, *762*(1-2), 29-39.

Smith, M. E., & Halgren, E. (1987). Event-related potentials during lexical decision: effects of repetition, word frequency, pronounceability, and concreteness. *Electroencephalography Clinical Neurophysiology Supplement*, *40*, 417-21.

Smith, M. E., Stapleton, J. M., & Halgren, E. (1986). Human medial temporal lobe potentials evoked in memory and language tasks. *Electroencephalography and Clinical Neurophysiology Supplement*, *63*(2), 145-59.

Solomon, E. S., & Pearlmutter, N. J. (submitted). Forward versus backward agreement processing in comprehension.

Solomon, E. S., & Pearlmutter, N. J. (2004). Semantic integration and syntactic planning in language production. *Cognitive Psychology*, *49*(1), 1-46.

Spitsyna, G., Warren, J. E., Scott, S. K., Turkheimer, F. E., & Wise, R. J. S. (2006). Converging Language Streams in the Human Temporal Lobe. *Journal of Neuroscience*, *26*(28), 7328.

Spitzer, M., Bellemann, M. E., Kammer, T., Gückel, F., Kischka, U., Maier, S., et al. (1996). Functional MR imaging of semantic information processing and learning-related effects using psychometrically controlled stimulation paradigms. *Cognitive brain research*, *4*(3), 149-161.

St. George, M., Mannes, S., & Hoffman, J. E. (1994). Global semantic expectancy and language comprehension. *Journal of Cognitive Neuroscience*, *6*, 70-83.

Stanovich, K. E., & West, R. F. (1981). The effect of sentence context on ongoing word recognition: Tests of a two-process theory. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(3), 658-672.

Stanovich, K. E., & West, R. F. (1983). On priming by a sentence context. *Journal of Experimental Psychology: General*, *112*(1), 1-36.

Staub, A., & Clifton Jr, C. (2006). Syntactic prediction in language comprehension: Evidence from either...or. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 425.

Stekelenburg, J. J., & Vroomen, J. (2007). Neural correlates of multisensory integration of ecologically valid audiovisual events. *Journal of Cognitive Neuroscience*, *19*(12), 1964-1973.

Stowe, L. A. (1986). Parsing WH-constructions: Evidence for on-line gap location. *Language and Cognitive Processes*, *1*(3), 227-245.

Stowe, L. A., Broere, C. A., Paans, A. M., Wijers, A. A., Mulder, G., Vaalburg, W., et al. (1998). Localizing components of a complex task: sentence processing and working memory. *Neuroreport*, *9*(13), 2995-9.

Stringaris, A. K., Medford, N. C., Giampietro, V., Brammer, M. J., & David, A. S. (2007). Deriving meaning: Distinct neural mechanisms for metaphoric, literal, and non-meaningful sentences. *Brain and Language*, *100*(2), 150-62.

Stroud, C. (2008). Structural and semantic selectivity in the electrophysiology of sentence comprehension. Ph.D. Dissertation, University of Maryland, College Park.

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, *48*(3), 542-562.

Suga, N., & Schlegel, P. (1972). Neural attenuation of responses to emitted sounds in echolocating rats. *Science*, *177*(43), 82.

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M. M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience*, *11*(9), 1004-1006.

Svoboda, E., McKinnon, M. C., & Levine, B. (2006). The functional neuroanatomy of autobiographical memory: a meta-analysis. *Neuropsychologia*, *44*(12), 2189-2208.

Swaab, T. Y., Brown, C., & Hagoort, P. (1997). Spoken sentence comprehension in aphasia: event-related potential evidence for a lexical integration deficit. *Journal of Cognitive Neuroscience*, *9*, 39-66.

Swick, D., Kutas, M., & Knight, R. T. (1998). Prefrontal lesions eliminate the LPC but do not affect the N400 during sentence reading. *Journal of Cognitive Neuroscience Supplement*, *29*.

Szewczyk, J. (2006). Anticipating animacy? An event-related brain potentials study of grammatic and semantic integration in Polish sentence reading. Paper presented at the Annual Meeting of the Society for Psychophysiological Research, Vancouver.

Tabor, W., Galantucci, B., & Richardson, D. (2004). Effects of merely local syntactic coherence on sentence processing. *Journal of Memory and Language*, *50*(4), 355-370.

Tanenhaus, M. K., & Lucas, M. M. (1987). Context effects in lexical processing. *Cognition*, *25*(1-2), 213.

Tarkiainen, A., Cornelissen, P. L., & Salmelin, R. (2002). Dynamics of visual feature analysis and object-level processing in face versus letter-string perception. *Brain*, *125*(5), 1125.

Tarkiainen, A., Helenius, P., Hansen, P. C., Cornelissen, P. L., & Salmelin, R. (1999). Dynamics of letter string perception in the human occipitotemporal cortex. *Brain*, *122*(11), 2119.

Thompson-Schill, S. L., Bedny, M., & Goldberg, R. F. (2005). The frontal lobes and the regulation of mental activity. *Current Opinion in Neurobiology*, *15*(2), 219-224.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proc Natl Acad Sci USA*, *94*(26), 14792-7.

Thompson-Schill, S. L., Kurtz, K. J., & Gabrieli, J. D. E. (1998). Effects of semantic and associative relatedness on automatic priming. *Journal of Memory and Language*, *38*(4), 440-458.

Thornton, R., & MacDonald, M. C. (2003). Plausibility and grammatical agreement. *Journal of Memory and Language*, *48*(4), 740-759.

Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage*, *39*(3), 1429-1443.

Traxler, M. J., & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language*, *35*(3), 454-475.

Tse, C. Y., Lee, C. L., Sullivan, J., Garnsey, S. M., Dell, G. S., Fabiani, M., et al. (2007). Imaging cortical dynamics of language processing with the event-related optical signal. *Proceedings of the National Academy of Sciences*, *104*(43), 17157.

Tyler, L. K., & Marslen-Wilson, W. (1986). The effects of context on the recognition of polymorphemic words. *Journal of Memory and Language*, *25*(6), 741-752.

Ullman, S. (1995). Sequence seeking and counter streams: a computational model for bidirectional information flow in the visual cortex. *Cerebral Cortex*, *5*(1), 1-11.

Uusvuori, J., Parviainen, T., Inkinen, M., & Salmelin, R. (2008). Spatiotemporal Interaction between Sound Form and Meaning during Spoken Word Perception. *Cerebral Cortex*, *18*(2), 456.

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating Upcoming Words in Discourse: Evidence From ERPs and Reading Times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(3), 443-467.

Van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: evidence from the N400. *Journal of Cognitive Neuroscience*, *11*(6), 657-71.

Van Berkum, J. J. A.The neuropgragmatics of 'simple' utterance comprehension: An ERP review. In U. Sauerland & K. Yatsushiro (Eds.), *Semantic and pragmatics: From experiment to theory*.

van den Brink, D., Brown, C. M., & Hagoort, P. (2001). Electrophysiological Evidence for Early Contextual Influences during Spoken-Word Recognition: N200 Versus N400 Effects. *Journal of Cognitive Neuroscience*, *13*(7), 967-985.

Van Gompel, R. P. G., & Liversedge, S. P. (2003). The influence of morphological information on cataphoric pronoun assignment. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *29*(1), 128-139.

Van Petten, C. (1993). A comparison of lexical and sentence-level context effects in event-related potentials. *Language and Cognitive Processes*, *8*, 485-531.

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition*, *18*(4), 380-93.

Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open- and closed-class words. *Memory & Cognition*, *19*(1), 95-112.

Van Petten, C., & Luka, B. J. (2006). Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain and Language*, *97*(3), 279-293.

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

Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 394-417.

Van Petten, C., Kutas, M., Kluender, R., Mitchiner, M., & McIsaac, H. (1991). Fractionating the word repetition effect with event-related potentials. *Journal of Cognitive Neuroscience*, *3*, 131-150.

Van Petten, C., Weckerly, J., McIsaac, H. K., & Kutas, M. (1997). Working memory capacity dissociates lexical and sentential context effects. *Psychological Science*, *8*(3), 238-242.

van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences USA*, *102*(4), 1181.

Vandenberghe, R., Nobre, A. C., & Price, C. J. (2002). The response of left temporal cortex to sentences. *Journal of Cognitive Neuroscience*, *14*(4), 550-60.

Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature*, *383*(6597), 254-6.

Vigliocco, G., & Nicol, J. (1998). Separating hierarchical relations and word order in language production: is proximity concord syntactic or linear? *Cognition*, *68*(1), 13-29.

Vissers, C. T., Chwilla, D. J., & Kolk, H. H. (2006). Monitoring in language perception: The effect of misspellings of words in highly constrained sentences. *Brain Research*, *1106*(1), 150-63.

Wagers, M. W. (2008). The Structure of Memory meets Memory for Structure in Linguistic Cognition. Ph.D. Dissertation, University of Maryland, College Park.

Wagers, M. W., & Phillips, C. (2009). Multiple dependencies and the role of the grammar in real-time comprehension. *Journal of Linguistics*, 45, 395-433.

Wagers, M. W., Lau, E., & Phillips, C. (2009). Agreement attraction in comprehension: representations and processes. *Journal of Memory and Language.*

Wagers, M. W., Lau, E., Stroud, C., McElree, B., & Phillips, C. (2009). Encoding syntactic predictions: evidence from the dynamics of agreement. Davis, CA.

Wagner, A. D., Koutstaal, W., Maril, A., Schacter, D. L., & Buckner, R. L. (2000). Task-specific repetition priming in left inferior prefrontal cortex. *Cerebral Cortex*, *10*(12), 1176-84.

Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning left prefrontal cortex guides controlled semantic retrieval. *Neuron*, *31*(2), 329-338.

Warren, R. M. (1970). Perceptual Restoration of Missing Speech Sounds. *Science*, *167*(3917), 392-393.

Wernicke, C. (1874). The aphasic symptom complex: A psychological study on a neurological basis. Breslau: Kohn and Weigert. Reprinted in *Boston studies in the philosophy of science*, *4*.

West, R. F., & Stanovich, K. E. (1978). Automatic Contextual Facilitation in Readers of Three Ages. *Child Development*.

Wheatley, T., Weisberg, J., Beauchamp, M. S., & Martin, A. (2005). Automatic priming of semantically related words reduces activity in the fusiform gyrus. *Journal of Cognitive Neuroscience*, *17*(12), 1871-85.

Wible, C. G., Han, S. D., Spencer, M. H., Kubicki, M., Niznikiewicz, M. H., Jolesz, F. A., et al. (2006). Connectivity among semantic associates: an fMRI study of semantic priming. *Brain and Language*, *97*(3), 294-305.

Wicha, N. Y. Y., Bates, E. A., Moreno, E. M., & Kutas, M. (2003). Potato not Pope: human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neuroscience letters*, *346*(3), 165-168.

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating Words and Their Gender: An Event-related Brain Potential Study of Semantic Integration, Gender Expectancy, and Gender Agreement in Spanish Sentence Reading. *Journal of Cognitive Neuroscience*, *16*(7), 1272-1288.

Wiggs, C. L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology*, *8*, 227-233.

Willems, R. M., Ozyurek, A., & Hagoort, P. (2008). Seeing and Hearing Meaning: ERP and fMRI Evidence of Word versus Picture Integration into a Sentence Context. *Journal of Cognitive Neuroscience*, *20*(7), 1235-1249.

Williams, G. B., Nestor, P. J., & Hodges, J. R. (2005). Neural correlates of semantic and behavioural deficits in frontotemporal dementia. *Neuroimage*, *24*(4), 1042-1051.

Wise, R. J., Scott, S. K., Blank, S. C., Mummery, C. J., Murphy, K., & Warburton, E. A. (2001). Separate neural subsystems within 'Wernicke's area'. *Brain*, *124*(Pt 1), 83-95.

Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, *3*, 1213.

Wright, B., & Garrett, M. (1984). Lexical decision in sentences: Effects of syntactic structure. *Memory & Cognition*, *12*(1), 31-45.

Xiang, M., Dillon, B., & Phillips, C. (2009). Illusory licensing effects across dependency types: ERP evidence. *Brain and Language*, *108*(1), 40-55.

Yoshida, M., & Sturt, P. (2009). Predicting 'or'. Poster presented at the 22nd Annual CUNY Conference on Human Sentence Processing, Davis, CA.

Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: analysis by synthesis? *Trends in Cognitive Sciences*, *10*(7), 301-308.

Zekveld, A. A., Heslenfeld, D. J., Festen, J. M., & Schoonhoven, R. (2006). Top–down and bottom–up processes in speech comprehension. *Neuroimage*, *32*(4), 1826-1836.

Zempleni, M. Z., Renken, R., Hoeks, J. C. J., Hoogduin, J. M., & Stowe, L. A. (2007). Semantic ambiguity processing in sentence context: Evidence from event-related fMRI. *Neuroimage*, *34*(3), 1270-1279.

Zipser, K., Lamme, V. A. F., & Schiller, P. H. (1996). Contextual modulation in primary visual cortex. *Journal of Neuroscience*, *16*(22), 7376-7389.