Predicting and Parsing Language in Time and Space

by

Suzanne Dikker

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Linguistics
New York University
September, 2010

________________________________________

Liina Pylkkänen – Advisor
DEDICATION

To you,

Joke, for thinking this owl a falcon
Loek, for pushing me into the deep end of the pool
Talia, for making sure I wouldn’t drown there
Piwi, Ariel, and Chuby, because you couldn’t care less
Damon, my ventilation system
Hugh, each and every one of you

Oh, and to New York City, of course
ACKNOWLEDGMENTS

Thank you, Liina Pylkkänen. I cannot imagine a more dedicated advisor. You pulled me along even if I was dragging my feet. This road would have been much longer and much less successful without you.

Thank you, neurolinguistics community. Everyone who stopped at posters, read and commented on manuscripts, and approached me out of the blue, making me feel like my research was worth doing. Thanks in particular to Gerry Altmann, for your helpful suggestions and ideas, to Nicola Molinaro, for your healthy dose of skepticism, and to Albert Kim, for repeatedly re-exciting me about what we do. Thank you, also, Maryland people. Colin Phillips, for always being interested and available. Ellen Lau, for the fruitful discussions about prediction. Masaya Yoshida, Brian Dillon (you should've come to NYU!), Matt Wagers, Diogo Almeida, and Jon Sprouse.

Thank you, NYU linguists and psychologists. Kyle Major, Cara Shousterman, and Jon Brennan. Eytan Zweig, Tuuli Adams, Jason Shaw, Kevin Roon, Oana Savescu Ciucivara, Amy Wong, Vincent Chanethom, Marcos Rohena-Madrazo, Kara Becker, Andrea Cattaneo, Thomas Leu, Lisa Levinson, Jesse Harris, Chris Barker, Alec Marantz, Gwyneth Lewis, Greg Cogan, David Poeppel. Andrew Smart; you brought some fresh air into that suffocating lab on the 7th floor with your weird sense of humor (:P). Jeffrey Walker, for always being there to help, for your amazing head digitizing skills, and for your patience. Yes, your patience. Brian McElree and Greg Murphy, for your open-door policy (and Brian, for keeping similar hours); Zvjezdana Vrzic and
Ofelia Garcia, for being such warm presences and for keeping me focused on the fact that it's all about being happy.

Thank you, Jon Brennan, for letting me provide you with an endless supply of vegan chocolate-chip cookies (yuk!) in return for your endless patience and help with anything, and pretty much everything, data-analysis related. And for your peer input on fears and doubts, and the laughs. Doug, for the same things, except the cookie part.

Thank you, all of you working in and around NYU, for your coffee and sandwiches, but most of all for life stories from Tunisia, Afghanistan, Morocco, Mexico, and the South-Bronx. Thank you, John Suffrand, for being the invisible elf who cleans our offices at 3am. Through my upside-down hours in the past few months, I finally appreciate and admire your work. A lot.

Thank you, Hakon Tretvoll, Peter Sokol-Hessner, and Hugh Rabagliati, for the 'weekly' breakfasts that forced me to wake up before 11am in spite of Hakon's eternal resistance. Robb Rutledge, Mat Coakley, Hakon Tretvoll, Peter-Sokol Hessner, Hugh Rabagliati, Andrea Martin, Jonny Cottrell, Lindsay Kaplan, Alex Irvine and Davita for that first year.

Thank you for living in NYC these past few years (and get back here asap!), Irene, Jake, Rudy, Grote Bas, Bitu, Julian, Elliott, James, and Pia. Talia and Marcelo, for taking in these sad, homeless, pieces of a girl, and providing such a warm nest during potentially the most stressful times of my life. Damon, for that one time in that one Mexican restaurant where you asked me “but where do these predictions come from?” right before I spilled all my food onto my lap after a sudden margherita-
induced clumsiness attack. And for listening to my rants and whines, over and over and over again.

Thank you, Amsterdam. Joke, Loek, Tessa, Jascha, Anjelene, Marianna, Frans, Nelleke, Matthias, oma Joyce, and Gert, for being such an amazing and supportive family. Kees Hengeveld, Dik Bakker, Ellen-Petra Kester, and especially Jacques Arends, for igniting my interest in Language. Jacques, you are missed. Eva van Lier and Josje Verhagen; Although we are now oceans apart, and it has been a while since we last sat down to discuss a bit of science and a lot of everything else, it still feels like we were in this together. Here's to you, from the last one to bite the dust.

Thank you for walking the face of the same planet as I do, Antonio, Sietske, Melle, Heleen, Marius, and Kleine Bas. Freddy, for being my best friend despite distance and years. Safe.

Finally: thank you, thank you, thank you, Hugh Rabagliati. For everything. For being the most amazing collaborator imaginable, for your patience, for reminding me that science CAN be really fun if it involves a whiteboard, a couch, and a 2-slice pizza deal. Thank you for keeping/making me sharp and for poking into the holes of my, um, 'associative' rants about the workings of prediction even long after you had a vested interest in the project. Thank you for being so annoying sometimes, for being giggly, for your ridiculous OCD screen-reading, for teasing Doug and Jon. For being an amazing friend, picking me up at times where I couldn't be more down. I will do miss you. Very, very much.
ABSTRACT

We use our experience to navigate the world in the present, and to predict what might come up in the near future. In this dissertation, I ask how prediction might help explain the rapidity and efficiency of language processing. In four experiments, I show, for the first time, that brain regions dedicated to low-level sensory analysis are sensitive to seemingly high-level factors of language processing as early as 100 ms after the presentation of a word. To explain this finding, I propose a Sensory Hypothesis: From contextual predictions, the brain derives form-feature estimates for syntactic categories or lexical-semantic representations, and these estimates are made available to sensory cortices by way of top-down modulation. A mismatch with a form expectation then causes enhanced neural activity in sensory regions.

While many scholars accept that prediction plays a crucial role in language processing, the neural correlates of linguistic prediction itself, or the preactivation of predicted linguistic representations, have not been previously investigated. I here compare highly predictive contexts to cases where no expectations are generated and show that left-temporal cortex (preactivation of lexical representations) and visual cortex (preactivation of form features) show increased brain activity for highly predictive environments right before the expected word is presented. The same brain areas were more active for words that violated lexical-semantic expectations than for words that satisfied predictions, but in reverse temporal order. As such, this research takes an important step towards elucidating the mechanisms by which prediction allows rapid language processing.

*Much of this research was conducted in collaboration with Hugh Rabagliati*
CONTENTS

DEDICATION .................................................................................................................. iv

ACKNOWLEDGMENTS .............................................................................................. v

ABSTRACT ................................................................................................................... viii

LIST OF FIGURES ...................................................................................................... xii

LIST OF TABLES ......................................................................................................... xiv

Chapter 1 - Introduction ............................................................................................ 1

  1.1 The syntax-first puzzle and prediction ............................................................... 2
  1.2 Part 1: Are early effects of linguistic violations sensory in nature? .................. 6
  1.3 Part 2: Can early sensory effects of linguistic violations be explained in terms
      of predictive processing? .................................................................................... 10

Chapter 2 - Early effects of linguistic violations: .............................................. 12

The sensory hypothesis .............................................................................................. 12

  2.1 A puzzle: the ELAN ............................................................................................. 12
  2.2 The Sensory Hypothesis .................................................................................... 15
  2.3 Potential generators of early effects .................................................................... 24
Chapter 3 - Experiments 1-4: Testing the Sensory Hypothesis ........... 27

3.1 Experiment 1: Testing the Sensory Hypothesis................................. 30
3.2 Experiment 2: Lexical access? Or closed-class morphemes only? .......... 35
3.3 Experiment 3: Closed-class morphology or form-typicality? ................. 41
3.4 Experiment 4: Also lexical-semantic violations? .................................. 47
3.5 Summary.......................................................................................... 51

Chapter 4 - Experiment 1 ......................................................................... 54

4.1 Introduction ...................................................................................... 54
4.2 Methods .......................................................................................... 55
4.3 Results ............................................................................................ 64
4.4 Discussion......................................................................................... 78

Chapter 5 - Experiment 2 ......................................................................... 80

5.1 Introduction ...................................................................................... 80
5.2 Methods .......................................................................................... 82
5.3 Results ............................................................................................ 87
5.4 Discussion......................................................................................... 91

Chapter 6 - Experiment 3 ......................................................................... 93

6.1 Introduction ...................................................................................... 93
6.2 Methods .......................................................................................... 96
6.3 Results ............................................................................................ 100
LIST OF FIGURES

Figure 1 - Research questions: predictive language processing in the brain .................... 11
Figure 2 - Hypothesis space for early effects of expectation violations ..................... 29
Figure 3 - Representational level accessible to visual cortex: Word category? ............. 38
Figure 4 - Representational level accessible to visual cortex: Form features?............... 43
Figure 5 - Summary of results for Experiments 1 - 4 ................................................. 53
Figure 6 - Grandaveraged waveforms and field distributions of the M100 (Exp. 1) ....... 65
Figure 7 - M100 dipole waveforms and peak amplitudes (Exp. 1) .............................. 67
Figure 8 - M170 dipole waveforms and peak amplitudes (Exp. 1) ............................... 70
Figure 9 - sensor waveforms and peak amplitudes in posterior octants (Exp. 1) ........... 74
Figure 10 - sensor waveforms and peak amplitudes in LA octants (Exp. 1) ................. 75
Figure 11 - M100 dipole waveforms and peak amplitudes (Exp. 2) ............................ 89
Figure 12 - M100 dipole waveforms per condition (Exp. 3) ..................................... 100
Figure 13 - M100 dipole waveforms for unambiguous nouns (Exp. 2) ......................... 105
Figure 14 - Stimulus presentation (Exp. 4) ......................................................... 120
Figure 15 - Grandaveraged field patterns for each component of interest (Exp. 4) ....... 127
Figure 16 - Trial-by-trial response times (Exp. 4) .................................................. 131
Figure 17 - M100 sensor waveforms per condition (Exp. 4) .................................... 132
Figure 18 - M170 sensor waveforms per condition (Exp. 4) ..................................... 135
Figure 19 - LA and RA sensor waveforms per condition (Exp. 4) ............................. 136
Figure 20 - Time-line stimulus presentation (Experiment 4) .................................. 150
Figure 21 - Predictions of competition models (Exp. 4) .............................................. 156
Figure 22 - Predictions of facilitation models (Exp. 4) .................................................. 159
Figure 23 - Prediction graphs for target word amplitudes (Exp. 4).............................. 162
Figure 24 - Competition vs. facilitation: M100 and left-anterior sensors (Exp. 4) ...... 165
Figure 25 - Competition mechanisms underlying M100 mismatch response (Exp. 4) 168
Figure 26 - Prestimulus activity; whole-brain minimum norm estimates (Exp. 4) .... 172
Figure 27 - Whole-brain minimum norm estimates for prestimulus and target word
brain activity at different frequency ranges (Exp. 4).................................................. 176
Figure 28 - Whole-brain minimum norm estimates for prestimulus and target word
brain activity in theta band (Exp. 4)............................................................................. 183
Figure 29 - Research questions: predictive language processing in the brain......... 196
LIST OF TABLES

Table 1 - Examples of experimental stimuli (Exp. 1) ................................................... 33
Table 2 - Examples of experimental stimuli (Exp. 2) ................................................... 36
Table 3 - Examples of experimental stimuli (Exp. 3) ................................................... 44
Table 4 - Examples of experimental stimuli (Exp. 4) ................................................... 49
Table 5 - Summary of hypotheses and results for Exps. 1-4.......................................... 52
Table 6 - Examples of experimental stimuli (Exp. 1) ................................................... 56
Table 7 - Examples of experimental stimuli (Exp. 2) ................................................... 84
Table 8 - Linear tegression results for single-trial M100 amplitude (Exp. 3) .............. 103
Table 9 - Examples of experimental stimuli (Exp. 4) .................................................. 1 20
Chapter 1 - Introduction

Language comprehension is restricted in a number of ways by external factors, the most obvious one being time: the speech signal reaches our auditory system sound-by-sound, and written language is typically read left-to-right. Although this is such a given that it is almost silly to even write it down, only very few linguistic theories take this linear sequencing into account (e.g., Steedman, 2000; Phillips, 2003; Kempson, Meyer-Viol, and Gabbay, 2001). Many either adhere to a “bottom-up” process of structure building, or do not consider directionality of structure-building altogether.

While detailed bottom-up syntactic theories as well as typological studies have proven very fruitful in the progress toward a comprehensive account of the nature of linguistic representations and how they combine, ultimately the inherently dynamic nature of language will have to be taken into consideration.

This dissertation asks how our parsing system deals with temporal dynamicity at the neural level, focusing on a subpart of the broad question of how brains process each incoming word most efficiently and rapidly: what is the role of prediction in explaining rapid effects of seemingly high-level factors of language processing?

The infinite possibilities of combining words into sentences make it impossible to always anticipate exactly what we will hear or read. On the other hand however, language is not a random scrambling of words but follows clear structural and semantic rules, and adult comprehenders have a vast experience with their own
language. Thus, although we can not always be 100% sure which word or morpheme will follow the one we just heard or read, structural and probabilistic regularities often allow us to make an informed guess.

This research asks when (time) and where (space) linguistic predictions affect neural responses. When and where are predictions generated, and what is the nature of the predicted representation? And when and where do we find effects of violations of predictions?

1.1 The syntax-first puzzle and prediction

Language is among the most complex of human cognitive systems, yet its processing is extremely automated and fast. Both behavioral and electrophysiological studies suggest that within 600ms of a word’s onset - whether presented auditorily or visually - its sensory properties have been analyzed, its syntactic and semantic features have been retrieved from memory, and it has been integrated into the preceding sentential context (e.g., Friederici, 2002; Kutas, Van Petten, and Kluender, 2006). A relatively consistent finding in electrophysiological studies of language processing is that certain syntactic factors affect neural processing at an earlier time-point than lexical-semantic violations: While event-related potentials (ERPs) to syntactic violations are observed as early as ~125ms after the onset of an unexpected word (the Early Left-Anterior
Negativity; ELAN; e.g., *The boys heard Joe’s about stories Africa*; e.g., Neville, Nicol, Barss, Forster, and Garrett, 1991; Friederici, Pfeifer, and Hahne, 1993), neural correlates of lexical, semantic, and world knowledge violations are typically indexed by a sustained negative-going wave peaking around 400ms (the N400 component; e.g., *I like my coffee with cream and socks*; Kutas and Hillyard, 1980; see Kutas et al., 2006 and Lau, Phillips, and Poeppel, 2008 for review). Some neuro-cognitive models of language processing have taken this pattern as evidence in favor of a syntax-first approach to language comprehension where the initial stage of processing reflects syntactic computation, only after which lexical and semantic factors come into play. Friederici’s (2002) model in particular is very explicit in assuming that different levels of analysis occur in a strictly autonomous and bottom-up fashion.

However, there exists increasing evidence that the brain is ‘proactive,’ continuously generating predictions about upcoming events. These predictions then affect processing of incoming stimuli. The role of predictive processing in language comprehension (see e.g., Federmeier, 2007; Pickering and Garrod, 2007; Levy, 2008 for recent accounts) has been demonstrated by cleverly designed experiments that show that predictions affect eye-movements and ERPs on words preceding an (un)predicted word (e.g., Altmann and Kamide, 1999; 2009; DeLong, Urbach, and Kutas, 2005; Van Berkum, Brown, Zwitserlood, Kooijman, and Hagoort, 2005; Wicha, Moreno, and Kutas, 2003; 2004).
More indirect evidence comes from effects on target words that are best explained in terms of prediction in reading time and eye-tracking experiments (e.g., McDonald, Scott, and Shillock, 2003; Staub and Clifton, 2006; Konieczny, 2000; Tanenhaus and Hare, 2007; Hale, 2001; Levy, 2008) as well as neuroimaging studies (e.g., Van Wassenhove, Grant, and Poeppel, 2005; Lau, Stroud, Plesch, and Phillips, 2006). For example, latency and amplitude variations of the N400 component appear to best explained in terms of facilitation of lexical access by way of sentence context or lexical priming (e.g., Federmeier and Kutas, 1999; Lau, Almeida, Hines and Poeppel 2009; see e.g., Lau, 2009 for review), rather than indexing ‘semantic integration’ (e.g., Hagoort, Bastiaansen, and Petterson, 2004). For syntactic violations, Lau et al. (2006) found a greater early negativity (ELAN) to violations of strongly, as compared to weakly, predicted word categories, and some studies have shown that the amplitude of early ERPs (in particular the P2 component) on target words is affected by whether or not sentential context was highly predictive (e.g., Wlotko and Federmeier, 2007).

These combined findings suggest that while the temporal sequence of ERP components related to different levels of language processing laid out in Friederici’s model (2002) might be largely empirically accurate (but see e.g., Hasting and Kotz, 2008), it is not very likely that the underlying theoretical assumptions of bottom-up processing can be held up.
However, accepting that previously generated predictions may affect processing, does not in itself explain the empirical generalization that syntactic effects take place before lexical-semantic effects.

One possibility that will be explored here is that there exists a difference between the detectability of violations of syntactic expectations and violations of lexical-semantic expectations: Maybe syntactic prediction facilitates processing at levels of representation that can be rapidly accessed after word presentation, while lexical-semantic prediction does not. More specifically, I will argue that while syntactic category predictions are translated into form-based estimates by way of top-down modulation of visual cortex, this is not always the case for lexical-semantic predictions.

This relates to the other sense of the “bottom-up vs. top-down” terminology: An increasing body of neurocognitive as well as anatomical evidence suggests that stimulus recognition relies on an interaction of different levels of representations (e.g., Lamme, 2003; Pammer et al., 2004; Bar et al., 2006; Yuille and Kersten, 2006; Rao and Hallard, 1999). For example, initial activation of visual cortex has been shown to trigger a rapid feed-forward sweep followed by slower-spreading activation both within higher-level processing areas and back to sensory regions (e.g., Lamme, 2003), challenging a purely bottom-up hierarchical view of the visual system. Pammer et al. (2004) suggest based on this type of evidence that as time progresses from stimulus-onset, ERP and MEG components are increasingly likely to index recurrent feedback-
driven, rather than bottom-up, processes. Bar et al. (2006), for instance, show that feedback-driven processes from Orbital-Frontal Cortex improve object recognition within 170 ms after stimulus onset, and demonstrations of top-down effects on letter and word recognition have been replicated across a range of experimental manipulations (e.g., Elman and McClelland, 1988; Duffy, Henderson, and Morris, 1989).

In sum, this dissertation explores the hypothesis that very early effects of linguistic violations are made possible through (i) predictive processing, where word categories are anticipated based on prior context, and (ii) top-down modulation of visual cortex, resulting from word category predictions being translated into form-estimates.

1.2 Part 1: Are early effects of linguistic violations sensory in nature?

In the first part of this dissertation (Chapters 2 through 7) I report on a series of findings that might begin to shed light on the syntax-first puzzle, by showing that syntactic violations affect the amplitude of very early responses generated in visual cortex, namely the visual M100 response as measured in magnetoencephalography (MEG) (Dikker, Rabagliati, and Pylkkänen, 2009; Dikker, Rabagliati, Farmer, and Pylkkänen, 2010).
As it has been shown that visual information may reach frontal regions as early as 80ms post-stimulus onset (Foxe and Simpson, 2002) and a feedback sweep back to visual cortex is be initiated within ~100ms post-stimulus onset (Lamme, 2003), it seems physically feasible that the M100 response (peaking around 100ms) could benefit from top-down modulation from higher-level language areas even in the absence of contextual prediction. However, this is unlikely given the absence of any evidence suggesting that the M100 component is sensitive to high-level linguistic factors for words in isolation (Tarkiainen et al., 1999; Solomyak and Marantz, 2009).

Instead, in Chapter 2, I argue in favor of a Sensory Hypothesis to explain early syntactic effects (Dikker et al., 2009). Under this hypothesis, context-based predictions about upcoming word categories include form-based estimates. These estimates are made available to sensory cortex by way of top-down modulation, as such giving rise to ‘syntactic’ effects during early visual analysis.

Experiments 1-4, summarized in Chapter 3, investigate the nature of the predicted representations to which the M100 response is sensitive.

While the results from Experiments 1 and 2 (Chapters 4 and 5; Dikker et al., 2009) suggest that word category effects observed during the M100 response might be restricted to closed-class category marking morphemes, which would be suggestive of visual cortex being exclusively sensitive to violations that are morpho-syntactic in nature, Experiment 3 (Chapter 6; Dikker et al., 2010) provides evidence that the visual M100 response is sensitive to the probabilistic distribution of form features across
syntactic categories (based on a phonological typicality measure conducted over English nouns and verbs by Farmer, Christiansen, and Monaghan, 2006). For example, in cases where a verbal form is predicted (e.g., the beautifully __), the amplitude of the M100 component in response to a noun like prince depends on how ‘nouney’ prince is: Higher M100 amplitude was found for nouns that have form features that are not shared with many verbs.

In sum, these findings suggest that the brain generates estimates about the likely physical appearance of upcoming words based on syntactic predictions: words that do not ‘look like’ the expected syntactic category show increased amplitude of the visual M100 component, the first salient MEG response to visual stimulation. This, in turn, might facilitate rapid successive syntactic structure building.

The observation that sensory cortices are sensitive to word category violations, a finding recently replicated for the auditory modality by Herrmann and colleagues (2009), may contribute to elucidating why syntactically relevant responses occur so early, but it still remains unclear why effects to lexical-semantic violations typically take place much later\(^1\). One possible explanation is that the absence of early visual effects for lexical-semantic violations is simply a byproduct of the nature of the linguistic system: While syntactic categories map onto certain feature combinations in a word’s form (e.g., Kelly 1992; Farmer et al. 2006), there is typically no

\(^1\)Although the term N400 is a bit misleading in this respect, as its onset is usually around 200ms post-stimulus onset, suggesting that lexically relevant processing begins at that latency or even earlier.
straightforward mapping between lexical-semantic fields (e.g., flowers) and visual or auditory forms (e.g., tulip, rose, magnolia). In other words: lexical-semantic properties do not translate into form-estimates as easily as syntactic categories do, at least in cases where more than a single lexical item can be anticipated.

Alternatively, sensory cortices may be exclusively tuned to detect syntactically relevant form-features. Or maybe predictions for syntactic category and lexical-semantic prediction entail qualitatively different processes and top-down modulation of visual cortex only occurs under syntactic anticipation. Neither of these hypotheses predicts that lexical-semantic violations would ever generate sensory effects.

However, under the hypothesis that linguistic predictions are translated into form-based expectations whenever possible, one would expect that lexical-semantic violations equally evoke a sensory effect under certain conditions. Experiment 4 (Chapter 7) indeed finds that the M100 response can be sensitive to violations of expectations for specific lexical items. This result suggests that syntactic features may not be ‘special’ during the earliest stages of analysis.

In sum, the combined finding from Experiments 1-4 suggest that early effects of linguistic violations can be explained in terms of violations of context-based form expectations, as such contributing to the increasing body of research that challenges a view of language processing whereby linguistic operations are strictly autonomous and take place in a serial and bottom-up fashion.
1.3 Part 2: Can early sensory effects of linguistic violations be explained in terms of predictive processing?

In the second part of this dissertation (Chapter 8) I take a stab at the elephant in the room: If it is true that the M100 effect falls out of form-expectations that are generated based on information analyzed at more ‘abstract’ levels of linguistic computation, then where and when are these expectations generated? In other words, what are the neural correlates of linguistic anticipation that make effects of violations of expectations possible? To my knowledge, these question have not been addressed before for sentence processing.

Experiment 4 begins to explore the neural correlates of linguistic prediction by investigating brain activity as a function of context, contrasting high cloze-probability (predictive) contexts with low cloze-probability (non-predictive) contexts. Above, I hypothesized that generating a prediction for a certain input should entail the preactivation of relevant representations in brain areas where said representations are stored, resulting in increased activity in these brain regions, as compared to cases where no prediction was generated. Thus, under the assumption that M100 effects reflect violations of word form predictions, one would expect to see more brain activity in visual cortex before the onset of (un)expected words in high-cloze probability contexts.
The results tentatively confirm these predictions. More activity in visual cortex was seen for highly predictive contexts compared to non-predictive contexts, right before stimulus-onset. In addition, immediately preceding this effect in visual cortex, a region in the left temporal lobe showed enhanced activation for predictive contexts, possibly reflecting preactivation of the lexical entry corresponding to the predicted word.

In sum, the research described in this dissertation is aimed at characterizing the neural correlates of prediction in language processing, investigating both the nature of the predicted representations and the anticipatory processes that make such expectations possible.

![Figure 1 - Research questions: predictive language processing in the brain](image)

*Experiments 1-4 investigate whether sensory cortices are sensitive to linguistic violations, and if so, what the nature is of the predicted representations that visual cortex operates on. Experiment 4 also asks which brain areas might be involved in predictive processing.*
Chapter 2 - Early effects of linguistic violations:  
The sensory hypothesis

2.1 A puzzle: the ELAN

One of the most striking findings from electroencephalography (EEG) is that syntactic operations can take effect as early as 120-160ms post-stimulus onset, at which point the presence of an unpredicted word category elicits an Early Left Anterior Negativity (ELAN) (Friederici et al., 1993; Neville et al., 1991). The rapidity of this response is surprising because the effect is cotemporaneous with early sensory processing (Bonte, Parviainen, Hytönen, and Salmelin, 2006; Di Russo, Martínez, Sereno, Pitzalis, and Hillyard, 2001; Hickok and Poeppel, 2007; Parviainen, Helenius, and Salmelin, 2005; Salmelin, 2007; Tarkiainen et al., 1999).

The ELAN refers to an enhanced negative-going wave generally found in left anterior electrodes in response to word category violations. This response has been reported for a number of languages, including English (Lau, Stroud, Plesch, and Phillips, 2006; Neville et al., 1991), German (Friederici et al., 1993; Hahne, Schroger, and Friederici, 2002; Rossi, Gugler, Hahne, and Friederici, 2005), Spanish (Hinojosa, Martin-Loeches, Casado, Muñoz, and Rubia, 2003), and French (Isel, Hahne, Maess, and Friederici, 2007). For example, Neville et al. (1991) reported an ELAN peaking
around 125ms post-stimulus onset for ungrammatically positioned prepositions like *about in sentence (1b).

(1)  a. The boys heard Joe’s stories about Africa.
    b. * The boys heard Joe’s about stories Africa.

In German, several studies have shown unexpected participles to generate an ELAN, for example when following the contracted preposition-determiner sequence im (2b), as opposed to after an inflected verb wurde (‘was’ 2a; taken from Friederici et al. (1993).

(2)  a. Das Baby wurde gefüttert
    The baby was fed
    b. * Das Baby wurde im gefüttert
    The baby was in the fed

Friederici (2002) proposes a model of the neurocognition of language that comprises a number of stages, each corresponding to a discretely defined electrophysiological correlate and brain area. Word category, morpho-syntax and semantics are autonomously, and not interactively, processed, and syntactic processing is initiated before semantic processing. In particular, the first stage of parsing a word in a
sentence is argued to be word category access (in the anterior superior temporal gyrus) and then, on the basis of this information, local phrase-structure is built (in the inferior portion of BA 44). Friederici suggests these processes are performed within 200ms of the word’s onset (Friederici, 2002; Friederici, Wang, Herrmann, Maess, and Oertel, 2000). In this model, the ELAN is the electrophysiological correlate of this early syntactic processing: an enhanced ELAN is generated by an ungrammatical word category, which impedes normal initial structure building. The ELAN therefore is argued to reflect the impossibility of reconciling the word category of the target with its local syntactic environment. In other words, the ELAN would be a response to grammatical ill-formedness rather than to a violation of a prediction. Consequently, this account thus assumes that a number of processes take place within ~120ms: (1) orthographic/phonetic processing; (2) word form access; (3) syntactic category retrieval; and (4) syntactic structure building.

This would require a surprisingly efficient and specialized processing system. Further, the functional interpretation of the ELAN as being restricted to local syntactic structure building is challenged by a recent finding that agreement violations generate an ELAN as well (Hasting and Kotz, 2008).
2.2 *The Sensory Hypothesis*

How is it physically feasible that seemingly high-level linguistic operations such as phrase-structure building can affect brain responses as early as 120ms after the presentation of a word? This question becomes increasingly puzzling when considering that the time window of the ELAN overlaps with early sensory processing (Bonte et al., 2006; Di Russo et al., 2001; Hickok and Poeppel, 2007; Parviainen et al., 2005; Salmelin, 2007; Tarkiainen et al., 1999).

One possibility is that the ELAN is a sensory response. Under this Sensory Hypothesis, linguistic predictions may include expectations that are made available to sensory cortices via top-down modulation. This in turn facilitates sensory effects of mismatches with such expectations.

To test this sensory hypothesis, a set of experiments were conducted using word-by-word reading in magnetoencephalography (MEG). MEG offers the best currently available tool for combining temporal and spatial resolution: Brain activity is monitored millisecond by millisecond and the current generators of the magnetic fields can be localized with relatively high accuracy (Hämäläinen et al., 1993). Specifically it was tested whether the earliest major visual response generated in MEG, the visual M100, is sensitive to linguistic manipulations, and, if so, under which conditions.
2.2.1 Properties of the ELAN that suggest its sensory nature

A number of properties of the ELAN are compatible with the hypothesis that the response might, at least in part, be generated in sensory cortices. These are discussed in turn below.

Latency

As discussed above, the first response to word category violations is cotemporaneous with early sensory responses. Both take place in the first 200ms post stimulus onset. In the auditory modality, early sensory activity is reflected by the N1 in EEG and the auditory M100 in MEG, both peaking at around 100ms. In the visual modality, sensory activity is measured by the P1/N1 complex in EEG (Di Russo et al., 2001; Luck, 2005) and the visual M100 in MEG, followed by the visual M170. The MEG M100 and the M170 have also been called “Type I” and “Type II” activity, respectively (Salmelin, 2007; Tarkiainen, Cornelissen, and Salmelin, 2002; Tarkiainen et al., 1999). Consistent with the timing of these various sensory components, the latency of the early responses to word category violations range from 125ms (Neville et al., 1991) to 160ms (Friederici et al., 1993; Friederici et al., 2000).

However, in certain circumstances, responses to word category violations have also been reported to begin later than the 100-200ms sensory time-window. In spoken
stimuli for example, there is some evidence that prefixed words generate an ELAN, but suffixed words generate a later response (typically referred to as a LAN), at 300-500ms (Friederici and Meyer, 2004). This, of course, is not surprising given that prefixes and suffixes become available at different time points in the spoken word. As Friederici (2001) discusses, the ELAN is obtained only after word category information has become available. Hence, an ELAN will be early for words containing prefixes, and later for words where the category information is contained in a suffix. Furthermore, some studies have indicated that visually presented suffixed words also yield a LAN as opposed to an early response, despite the immediate availability of the suffix in the visual modality. Importantly, though, most of these manipulations have involved inflection violations rather than phrase-structure violations (e.g., Barber and Carreiras, 2005; Morris and Holcomb, 2005; Rodríguez-Fornells, Clahsen, Lleo, Zaake, and Munte, 2001; Rossi et al., 2005), which arguably involve different processes. There is one study that reports a LAN instead of an ELAN for word category violations (Hagoort, Brown and Wassenaar, 2003). However, in this study the word category violations were indirectly marked through an agreement morpheme (3rd person singular verbal morphology on an ambiguous root).

In sum, in the auditory modality, early responses to word category violations are consistently found in the same time period as sensory processing, i.e., within 200ms after the word category information becomes available. In the visual modality, however, the corresponding generalization is somewhat less robust.
Localization

In Friederici’s (2002) model, the ELAN is generated by sources in the left inferior frontal gyrus (IFG) and superior temporal gyrus (STG). These localizations are based on evidence from fMRI, lesion, and MEG studies.

In fMRI, Meyer, Friederici, and Von Cramon (2000) used auditorily presented stimuli involving a participial manipulation similar to the one already exemplified in (2) above. Ungrammatical, as compared to grammatical, stimuli resulted in a reliably enhanced BOLD signal in the anterior portion of the superior temporal gyrus. Importantly however, the BOLD response was also enhanced in both primary and secondary auditory cortex (including Heschl’s gyrus and the planum temporale) – sensory cortices typically associated with speech processing and the generation of auditory evoked responses (Hickok and Poeppel, 2007; Näätänen, 1992). Enhanced activity in auditory cortices was also found in a second study, using similar stimuli, which replicated Meyer et al.’s results in primary and secondary auditory cortex (Friederici, Rüschemeyer, Hahne, and Fiebach, 2003), and the anterior superior temporal gyrus. Together, these findings suggest that sensory cortices, as well as left anterior temporal regions, are sensitive to word category violations. However, they do not address which regions are active during the ELAN time window, given the low temporal resolution of fMRI.

In lesion studies, the absence of an ELAN in Broca’s aphasics with damage to the left inferior frontal gyrus (IFG) (Friederici, Von Cramon, and Kotz, 1999) has
been taken as evidence that early responses to word category violations are generated in the left IFG. However, the hypothesis that this area is part of a cortical network that generates the predictions for upcoming word categories is equally compatible with the data. In other words, if the left IFG is involved in generating word category predictions rather than checking whether the input matches them, then damage to this area would block the generation of predictions. Consequently, there would be no predictions for the input to violate and no ELAN would be observed.

Thus both fMRI and lesion studies underdetermine the neural generators of early responses to word category violations. Magnetoencephalography (MEG) offers the requisite temporal and spatial resolution to potentially address this question: brain activity can be monitored millisecond by millisecond and the current generators of the magnetic fields can be localized with relatively high accuracy (Hämäläinen, Hari, Ilmoniemi, Knuutila, and Lounasmaa, 1993). In an auditory MEG study, Friederici et al. (2000) presented subjects with grammatical and ungrammatical participles similar to example (2) above. A four-dipole model was fit to the ungrammatical data only, including sources in the left and right anterior portions of the superior temporal gyri as well as in fronto-lateral cortices bilaterally. From the model’s relatively high goodness-of-fit, it was concluded that the ELAN is generated by these four sources. However, the modeling procedure used in this study makes this conclusion somewhat premature. To create each subject’s source-model, four dipoles were seeded in bilateral frontal and temporal locations. Each dipole’s orientation and location were fit
to the violation conditions’ data within a 20ms interval centered on the peak of the subject’s ELAN response. Variation in location, however, was constrained to a sphere with a 10mm radius, centered on the initial seed point. Since this constraint \( a \ priori \) ensured localization of the ELAN close to the seeded locations, it is difficult to draw firm conclusions, despite the model’s relatively high goodness-of-fit. Interestingly, for all subjects the dipole strength of the temporal source in the 20ms time window was higher than that of the frontal source. Given the greater proximity of the temporal (as opposed to frontal) dipole to auditory cortex, this is consistent with enhanced sensory activity, as predicted by the hypothesis that sensory cortices contribute to early effects of word category violations. Perhaps clearer evidence for the sensory hypothesis was obtained in an auditory MEG study by Gross et al. (1998), who reported enhanced and delayed activity in both the auditory cortex and left frontal cortex for word category violations, using Magnetic Field Tomography, a distributed source model. However, the authors did not expand on the possible implications of this finding.

The results of a recent EEG study, using visual stimuli, are also suggestive of the contribution of sensory cortices to the detection of word category violations. Specifically, Hagoort, Wassenaar, and Brown (2003) reported a significant interaction of grammaticality by electrode site at 100-300ms for Dutch word category violations. The interaction was due to enhanced activity in posterior electrodes (i.e., over the visual cortex), which in isolation however did not show a reliable modulation.
Further properties of the ELAN consistent with a sensory generator

If the sensory cortices are sensitive to word category violations, the topography of the ELAN might plausibly resemble that of sensory responses. This prediction is supported by findings from Hahne et al. (2002), who reported that the topography of early responses to word category violations was statistically indistinguishable from responses to deviance in low-level auditory stimulation. In this study, word category violations of the type in (2) above were crossed with a manipulation of auditory expectancy, where deviants were presented from a different location from the rest of the sentence fragment. This physical manipulation was predicted to elicit an auditory Mismatch Negativity (MMN) (Näätänen, 1992). The primary generators of the auditory MMN have been reported as the planum temporale (Liegeloi-Chauvel, Musolino, Badier, Marquis, and Chauvel, 1994), and near-by regions around auditory cortex (Virtanen, Ahveninen, Ilmoniemi, Näätänen, and Pekkonen, 1998). In contrast, as pointed out above, in Friederici’s model the ELAN is generated by two sources in the anterior left superior temporal gyrus, and in the inferior portion of BA 44. Instead of separate responses for the two manipulations, Hahne et al. (2002) found no topographical differences between the word category deviances and the physical deviances, consistent with the sensory hypothesis outlined above.

Similar results have also been obtained by Pulvermüller and colleagues with an oddball paradigm using linguistic stimuli only. Using MEG, Shtyrov, Pulvermüller, Näätänen, and Ilmoniemi (2003) found that hearing deviant syntactic violations
(16.7% of the stimuli) interspersed with standard grammatical sequences (83.3% of the stimuli) resulted in magnetic mismatch negativity (MMNm), the MEG response corresponding to the classical MMN (Näätänen, 1992). Minimum-norm estimation suggested the effect was generated in the left superior temporal lobe (although see Pulvermüller and Shtyrov (2003) for evidence from EEG for an additional frontal generator). Pulvermüller and Shtyrov (2006) argued that the syntactic MMN is highly similar to the ELAN in terms of latency, topography, and laterality. Although these authors did not contrast linguistic and non-linguistic manipulations directly, this result in combination with Hahne et al.’s (2002) findings suggests that early responses to word category violations may be generated by the same sources that respond to low-level auditory deviance.

Finally, there is also evidence in the visual modality that early syntactic responses may be modulated by the same factors that affect sensory components. Gunter, Friederici, and Hahne, (1999) demonstrated that no ELAN is observed in the visual modality when a target word is presented against a low contrast (dark grey) background. This result was interpreted to indicate that fast and automatic structure building is dependent upon good signal quality. However, this finding is equally consistent with the sensory hypothesis. In other words, low signal quality could hamper the sensory detection of a mismatch between word category expectations and properties of the visual (or auditory) input.
In sum, the ELAN literature at large is quite consistent with the hypothesis that sensory cortices are sensitive to word category violations. However, the picture is far from conclusive. FMRI provides no information regarding the timing of the effects in auditory cortex and EEG may yield null results regarding the topography of the ELAN vs. the MMN because of its low spatial resolution.

To address the sensory hypothesis we used MEG, which offers the same temporal accuracy as EEG but with enhanced spatial resolution. Further, we presented stimuli visually, as opposed to the auditory presentation used in the previous MEG studies reviewed above (Friederici et al., 2000; Gross et al., 1998). An obvious implication of the sensory hypothesis is that responses to word category violations should be modality dependent: one would expect modulation of activity in auditory cortex when stimuli are presented aurally and in visual cortex when stimuli are presented in the visual modality. We chose the visual modality since the visual cortex is further away from left anterior regions than the auditory cortex. Thus detecting a sensory modulation not attributable to left anterior generators should be easier in the visual than in the auditory modality. The next section reviews the visual evoked components relevant to the present experiments.
2.3 Potential generators of early effects:

The M100 and M170 components

Early stages of visual word recognition in MEG are primarily characterized by two response components. The first component is the co-called visual M100, also called “Type I activity” (Tarkiainen et al., 2002; Tarkiainen et al., 1999). This response is generated bilaterally in the occipital lobe, close to midline, at 100-150ms, i.e., in a similar time window as the ELAN. A detailed localization study using the Synthetic Aperture Magnetometry technique has suggested that the visual M100 has maximum intensity in the cuneus, lingual gyrus, and BA 17 (Itier, Herdman, George, Cheyne, and Taylor, 2006). Due to the close proximity of the bilateral generators, M100 activity is usually modeled by a single dipole, although bilateral two-dipole solutions are also possible (e.g., Pylkkänen, Llinas, and Murphy, 2006). The field distribution of the M100 typically shows a single posterior right-lateralized outgoing magnetic field and a left-lateralized posterior re-entering field (for examples, see Itier et al., 2006; Pylkkänen et al., 2006). Previous studies on words in isolation strongly suggest that the visual M100 is a pre-lexical, low-level visual response. For example, Tarkiainen et al. (1999) did not report any differential activity at the M100 response to letter strings compared to symbols. Similarly, in a recent MEG study using a lexical decision task, no effects of lexical factors were found before 150ms (Solomyak and Marantz, 2009).
An exception to this has been observed in the domain of face perception, where the visual M100 has been shown to be modulated by face categorization (Liu, Harris, and Kanwisher, 2002).

The second prominent response to visual stimulation is the M170 component, or “Type II activity” (Tarkiainen et al., 2002; Tarkiainen et al., 1999), peaking at 150-200ms. fMRI evidence from manipulations similar to those affecting the M170 suggests that the M170 is generated in the left and right fusiform gyri (Cohen et al., 2000; Dehaene et al., 2004; Kanwisher, McDermott, and Chun, 1997; Tong, Nakayama, Vaughan, and Kanwisher, 1998). The left and right generators of the M170 appear to differ in functional specialization: the left generator has been reported as sensitive to the presence of letter-strings (Tarkiainen et al., 2002; Tarkiainen et al., 1999), whereas the right generator shows increased amplitudes for faces in comparison to various control categories (e.g., Kanwisher et al., 1997; Liu et al., 2002; Lueschow et al., 2004; Tarkiainen et al., 2002; although see Liu, Higuchi, Marantz, and Kanwisher, 2000, for a bilateral M170 modulation for face stimuli). In the fMRI literature, the left and right fusiform regions sensitive to letter-strings and faces have been dubbed the Visual Word Form Area (VWFA) and the Fusiform Face Area (FFA) respectively. In addition to showing sensitivity to the linguistic vs. non-linguistic nature of stimuli, the M170 has recently been reported as sensitive to morphological complexity, showing higher amplitudes for bimorphemic than for orthographically matched monomorphemic words (Zweig and Pylkkänen, 2009).
On the sensory hypothesis, both the M100 and the M170 components constitute candidates for an early visual effect of word category expectations. Given that the M170 has already been reported to be sensitive to linguistic factors, an M170 effect of category mismatch would perhaps be less surprising than an M100 modulation. However, the latency of the M100, which on average peaks at 130ms, is in fact in better correspondence with ELAN latencies than the M170. In studies using English stimuli, the ELAN has peaked exactly around 130ms (Neville et al., 1991; Yamada and Neville, 2007). The M100 is also sensitive to the same noise manipulations as the ELAN (Gunter et al., 1999; Tarkiainen et al., 2002).

It is of course also possible that the M100 and the M170 both show sensitivity to word category violations. Such a result would in fact conform well with the fMRI finding that in the auditory modality, word category violations elicit an increased BOLD signal across all areas of the superior temporal gyrus. However, as will become clear below, the M100 component alone, and not the M170 component, appears to be sensitivity to word category violations.
Chapter 3 - Experiments 1-4: Testing the Sensory Hypothesis

In what follows, I summarize the results of four experiments testing the Sensory Hypothesis. Full descriptions of each experiment can be found in Chapters 4-7. Figure 2 presents an overview of the main research questions and corresponding hypotheses involving the possible levels of representation available to the M100 response – ranging from lexical access to visual form analysis. In Experiment 1 the hypothesis is tested that the visual M100 response, the first salient MEG response to visual stimulation, is sensitive to word category violations.

Experiments 2-4 then examine the conditions under which an M100 effect is observed: Experiment 2 asks whether nouns which are unambiguous for word category but do not contain any word category markers, like closed-class morphemes, generate an M100 effect in unexpected contexts. If so, this might indicate that visual cortex has access to some form of lexical representation. If not, the question remains if early sensory effects of word category violations might reflect a mismatch at the level of word category analysis. This is related to the question of whether closed-class morphology is necessary: Experiment 1 and 2 found effects for target words that contained a closed-class category marking morpheme, but crucially not for words without such elements. As will be explained below, under the hypothesis that closed-
class morphology is necessary, this pattern of results leads to the conclusion that the M100 generator supports word category analysis.

However, if closed-class morphology is not necessary to generate an M100 effect for unexpected word categories, but general form features that are indicative of word category suffice, then one need not conclude that the M100 generator supports word category analysis. This is tested in Experiment 3. Experiment 4 then asks if the M100 generator might be sensitive to violations of word form predictions in general, whether based on syntactic or lexical-semantic predictions.
Figure 2 - Hypothesis space for early effects of expectation violations

Experiment 1 tests whether visual cortex is sensitive to syntactic violation. If so, the visual M100 (or M170) component should be sensitive to word category violations (WC). Experiment 2 explores if the M100 generator supports lexical access. If not, WC unambiguous nouns should not generate an M100 effect in unexpected contexts in the absence of closed-class category marking morphology (CCM; e.g., tree as opposed to princess). Experiment 3 then tests whether the M100 generator supports syntactic category (WC) analysis. If not, words that violate a form expectation but do not themselves contain a CCM should also generate an M100 effect in unexpected contexts. Finally, Experiment 4 investigates whether the M100 generator is sensitive only to syntactic violations, or also lexical-semantic violations.
3.1 Experiment 1: Testing the sensory hypothesis

See Chapter 4 for a full description of Experiment 1.

3.1.1 Introduction and Design

In Experiment 1, we compared different types of word category violations that have previously generated an ELAN response. The first comparison consisted of the same set of stimuli tested by Neville et al. (1991), as exemplified in (1) (*The boys heard Joe's stories about Africa* vs. *The boys heard Joe's about stories Africa*)

The second manipulation involved participles. This manipulation is an adaptation into English of the German phrase-structure violations exemplified in (2) (e.g., Friederici et al., 1993), where a participle followed either a copula (expected), or a determiner (unexpected) (*The discovery was reported* vs. *The discovery was in the reported*). Note that a participial modifier of a yet-to-come noun is perfectly grammatical in this position. However, a participle is not the most likely continuation of a determiner: In an off-line sentence completion task, participants (N = 60) inserted a noun after ‘the’ in 91% of the sentence frames, while only very few filled in a participle.

Both prepositions and participles contain closed-class morphemes. Many behavioral studies have suggested that closed-class morphemes are processed differently from open-class items. In particular, Bradley (1983) has argued that the
lexical retrieval of closed-class morphemes is faster than that of open-class items and that closed-class items do not exhibit frequency effects (although see, e.g., Gordon and Caramazza, 1985; Segalowitz and Lane, 2000). There is also evidence that subjects are faster to reject non-words based upon closed-class items to which suffixes have been appended (e.g., thanage), than to equivalent open-class items (e.g., thinage) (Bradley, 1983; Matthei and Kean, 1989). In addition, eye-tracking studies have shown that while readers tend to fixate open-class words, they frequently skip function words (Carpenter and Just, 1983; Liversedge and Findlay, 2000; Reichle, Rayner, and Pollatsek, 2003). Thus the access of closed-class items appears to be extremely quick and automatic. Consequently, the rapidity of word category effects may in part be due to the rapid and automatic retrieval of frequent word category marking function morphemes.

A possible connection between the ELAN and closed-class morphology has been noted before. In a 1999 paper, Hahne and Friederici observed that thus far the ELAN had only been generated by unexpected or ungrammatical target words containing highly frequent, closed-class morphemes (Hahne and Friederici, 1999). To the best of my knowledge, this fact still holds. Several studies have found an ELAN for ungrammatical or unexpected verbal morphology in contexts where a noun or other word category was expected. ELAN responses for these types of word category violations have been found in German (Friederici, 2002; Friederici et al., 1993; Hahne and Friederici, 1999; Hahne et al., 2002), (for examples of these see (2) above);
(Oberecker, Friedrich, and Friederici, 2005), French (Isel et al., 2007), and Spanish (Hinojosa et al., 2003). In English, an early response has been observed in response to unexpected prepositions (Lau et al., 2006; Neville et al., 1991; see (1) above). In one study, in contrast, where word category was violated but the target item did not contain a closed-class morpheme, no ELAN was generated (Friederici and Meyer, 2004).

In sum, the access of closed-class items appears to be extremely quick and automatic. In view of this, the observation that early responses have only been reported for word category violations involving overt closed-class items suggests that their rapidity may at least in part be due to the rapid and automatic retrieval of frequent word category marking function morphemes. Thus sensory cortices might be tuned to recognize highly frequent, salient function morphemes that mark word category information, and show more activity when the input contains unexpected word category marking morphemes.

To test whether an M100 effect is contingent upon the present of an overt word category marker, we thus included a third comparison containing bare stems (*The discovery was in the report* vs. *The discovery was report*), i.e., lacking closed-class morphology.
Table 1 - Examples of experimental stimuli (Experiment 1)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Expected Example sentence</th>
<th>Cloze-P</th>
<th>Unexpected Example sentence</th>
<th>Cloze-P</th>
<th>Nature of the violation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preposition</td>
<td>The boys heard Joe's stories about Africa.</td>
<td>0.441</td>
<td>The boys heard Joe's about</td>
<td>0.000</td>
<td>Presence of an unpredicted free closed-class morpheme.</td>
</tr>
<tr>
<td>Participle</td>
<td>The discovery was reported</td>
<td>0.287</td>
<td>The discovery was in the reported.</td>
<td>0.002</td>
<td>Presence of an unpredicted bound closed-class morpheme.</td>
</tr>
<tr>
<td>Bare stem</td>
<td>The discovery was in the report.</td>
<td>0.900</td>
<td>The discovery was report.</td>
<td>0.002</td>
<td>Absence of a necessary bound closed-class morpheme.</td>
</tr>
</tbody>
</table>

Examples of the experimental stimuli of Experiment 1. 300 target sentences (60 per condition for Participles and Bare stems, 30 per condition for Prepositions) were intermixed with 240 filler sentences. Cloze-probabilities reflect the proportion of the word category of interest filled in during an off-line sentence completion task (see text): cloze-probabilities for unexpected targets were at or close to zero.

Paradigm/Task

Participants performed sensicality judgments on sentences such as the ones exemplified in Table 1, presented word-by-word (300ms on/off).
3.1.2 Results

Results revealed an effect of expectedness on M100 amplitude for prepositions and participles, i.e., those manipulations that have previously shown an ELAN effect. There was no clear additional left-anterior source responsive to phrase-structure violations (See Figures in Chapter 4). This finding is strongly suggestive of the sensory nature of the ELAN, although it is possible that an additional left-anterior effect was not picked up in our data. Regardless of this possibility, these results support the sensory hypothesis at least insofar as that they confirm that sensory responses are sensitive to syntactic manipulations.

In contrast to the prepositions and participles, the bare stem manipulation did not generate an M100 effect. This is suggestive of a special status of closed-class morphemes. However, the results of experiment 1 do not necessarily imply that closed-class morphemes have a special status.

First, it is possible that unexpected bare stems failed to generate an M100 effect because they were all category-ambiguous between nouns and verbs (e.g., report). This possibility is tested in Experiment 2.

Second, closed-class morphemes are not the only source of overt word category information in English: It has been shown that certain phonological feature-combinations occur more in verbs than nouns and vice versa. Experiment 3 tests whether the M100 generator is sensitive to such probabilistic indices of word category.
3.2  **Experiment 2: Lexical access? Or closed-class morphemes only?**

See Chapter 5 for a full description of Experiment 2.

3.2.1  **Introduction and Design**

Experiment 2 was designed to assess whether the bare stems of Experiment 1 failed to show an M100 effect because of their category ambiguity or because of a lack of closed-class morphology. To test this, a manipulation with unambiguously nominal stems like *tree* was included in addition to the participles and the ambiguous stems tested in Experiment 1 (e.g., *report/reported*). Further, predictability for a participle was maximized by inserting an adverb after the copula (raising the cloze-probability for a participle from 30% to 80%, see Chapter 5). The design is summarized in Table 2.
Table 2 - Examples of experimental stimuli (Experiment 2)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Expected</th>
<th>Unexpected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Category</td>
<td>Category Ambiguity</td>
<td>Example sentence</td>
</tr>
<tr>
<td>participle</td>
<td>Unambiguous - CCM</td>
<td>The discovery was solemnly reported.</td>
</tr>
<tr>
<td>bare stem</td>
<td>Ambiguous - CCM</td>
<td>The discovery was in the solemn report.</td>
</tr>
<tr>
<td>Noun</td>
<td>Unambiguous - CCM</td>
<td>The owl was in the high tree.</td>
</tr>
</tbody>
</table>

Examples of the experimental stimuli of Experiment 2 (60 per condition). Cloze-probabilities reflect the proportion of the word category of interest filled in during an off-line sentence completion task (see text). For the unexpected conditions, cloze-probabilities are also given for the violated word category expectation (i.e., participles for the unexpected bare stems and the unexpected unambiguous nouns, and nouns for the unexpected participles).

Paradigm/Task

Participants performed sensicality judgments on sentences such as the ones exemplified in Table 1, presented word-by-word (300ms on/off).

3.2.2 Results

There was an M100 effect for participles but not for bare stems, replicating the results of Experiment 1. Crucially, unexpected unambiguous nouns did not generate an M100 effect. This indicates that no category-disambiguation at the lexical level takes
place during the M100 component, supporting previous findings that the M100 response is pre-lexical (e.g., Solomyak and Marantz, 2009).

3.0  Word category analysis or form prediction mismatch?

The findings of Experiments 1 and 2 are suggestive of the neural reality of the distinction between open-class and closed-class elements. This raises the question what exactly might be special about closed-class category marking morphemes: Are they rapidly analyzed as word category markers during early visual responses? Or is the M100 effect a result of a mismatch between the expected and unexpected form of a CCM? Figure 3 lays out the mechanisms that would be at play for each of these two alternative hypotheses.
Figure 3 - Representational level accessible to visual cortex: Word category?

Two accounts of the representational level of closed-class morphemes (CCMs) at which a prediction mismatch is detected during the M100 response for each of the sentence contexts of Experiments 1 and 2: mismatch with predicted CCM form (A and B); or a mismatch with the word category (WC) indexed by the CCM (C and D)
Panels A and B illustrate that if an M100 effect were the result of a mismatch with the predicted form of a closed-class category marking morpheme (CCM), we would predict an effect for unexpected bare stems (A), but not for unexpected participles (B). A sentence frame like *The discovery was* __ (A; left) triggers an expectation for a participle, i.e., a word containing the closed-class category marking morpheme –ed. In the bare stem condition then, this expectation is violated, which should generate an M100 effect. *The discovery was in the* __ (B; left), on the other hand, triggers an expectation for a noun, which presumably does not include a prediction for any specific closed-class morpheme, as nouns in English often lack such category marking. Thus, for the conditions exemplified in A and B we would expect an M100 effect for cases where a prediction for a visual cue is violated (the bare stem manipulation). Instead, we find an effect for the participle manipulation, where there was no strong prediction for a closed-class morpheme.

Panels C and D exemplify the alternative possibility, where word category predictions are made available to visual cortex: *The discovery was* __ (C, left) triggers the expectation for a participle; for *The discovery was in the* __ (D, left) an expectation for a noun is generated. During or before the M100 response then, the input is analyzed for word category. In cases where the input contains a bare stem (instead of a participle; C, right) the absence of a closed-class category marking morpheme precludes word category analysis, and the mismatch cannot be detected in visual cortex. In contrast, if the target word contains the closed-class category marking
morpheme [-ed] (D, right) in conditions where a noun was expected, word category analysis is possible, and the mismatch can be detected.

In sum, the results presented so far suggest that the M100 generator must ‘know’ that –ed is a participial suffix, rather than blindly checking form predictions for closed-class category marking morphemes against the input. I.e., we are forced to conclude that the representations of closed-class morphemes available during the M100 response include word category information.
3.3 **Experiment 3: Closed-class morphology or form-typicality?**

See Chapter 6 for a full description of Experiment 3.

3.3.1 **Introduction and Design**

In the previous section, I explained how the combined results from Experiments 1 and 2 may indicate that the M100 generator supports word category analysis. This is highly surprising in view of extant evidence suggesting that the M100 response is sensitive to low-level visual properties only (Tarkiainen et al., 1999). However, while Experiments 1 and 2 demonstrate that it is unlikely that lexical access occurs during the M100 response, they do not necessarily prove that closed-class morphology is the driving factor. It is well-known that a number of other word form properties are indicative of word category, such as stress patterns and phonological combinations (e.g., Kelly, 1992). Thus it entirely possible that an M100 effect was obtained because words like *reported* are not probable nominal forms – i.e., nouns do not usually end in *–ed* – and that it is this anti-canonicity rather than the presence of a linguistically relevant morpheme /ed/ that drives the effect. Experiment 3 tested this possibility by manipulating form typicality associated with word category membership.

This form-typicality hypothesis derives from research demonstrating that systematic, probabilistic, form-based regularities exist among the words of a given lexical category, and that these regularities have consequences for on-line syntactic
processing (Arciuli and Monaghan, 2009; Cassidy and Kelly, 1991; Farmer et al., 2006; Fitneva et al., 2009; Kelly, 1992; Monaghan et al., 2007). In one recent study, Farmer et al. (2006) demonstrated via a corpus analysis that English nouns and verbs form clusters in phonological space, reflecting the relative occurrence of certain features in either category. While most nouns and verbs are ‘typicality neutral,’ containing form features that are equally common in both categories, there are also clearly typical nouns and verbs (more ‘typical’ nouns share less features with verbs and vice versa). Farmer et al. (2006) found that English speakers were faster to read these. Based on this work, Tanenhaus and Hare (2007) argued that form typicality may influence the earliest stages of fixation patterns during word reading. This would be consistent with an early visual M100 effect. Crucially then, the visual M100 component might be sensitive to the probabilistic distribution of form features across the entire mental lexicon, instead of being specifically tuned to detecting a small set of closed-class morphemes. M100 amplitude might be a function of the extent to which such form-based expectations are met. This hypothesis is exemplified in Figure 4. A prediction for a participle, for example, would be translated into an expectation of the form properties associated with participles. In Figure 4, this estimate is exemplified as the form typicality score [.2] (see Farmer et al. 2006; Dikker et al., 2010; Chapter 6 of this dissertation). During presentation of a noun then, M100 amplitude depends on how ‘close’ this particular noun’s typicality score is to [.2]: Higher M100 amplitude
would be expected for nouns containing features that lie farther on the typicality scale (e.g., [-.4] as opposed to [0])

**Figure 4 - Representational level accessible to visual cortex: Form features?**

*Illustration of the M100 response to word category violations stated in terms of a form-feature mismatch account, where the representational level is form typicality (see text).*

To test whether closed-class morphemes have a special status as category indicators, or whether form typicality can also serve as a category cue for visual cortex, we examined the visual M100 effect for three types of nouns presented in expected or unexpected contexts: (i) bimorphemic nouns (with a closed-class category marking suffix like *farm-er/princ-ess, art-ist*); (ii) monomorphemic ‘typical’ nouns containing form properties that are indicative of the noun category (e.g., *movie, soda*; mean typicality score around [-.2]), and (iii) neutral nouns (no clear form bias toward either nouns or verbs; typicality score around [0]). Bimorphemic nouns and typical nouns were about equally indicative of the noun category. To manipulate syntactic context, the critical noun was preceded by either an adjective (*the beautiful__*), where a noun
is highly expected, or by an adverb (the beautifully___), rendering nouns unexpected and instead inducing a strong expectation for a participle (like dressed). The design is exemplified in Table 3 (see Chapter 6).

Table 3 - Examples of experimental stimuli (Experiment 3)

<table>
<thead>
<tr>
<th>Noun type</th>
<th>Expected</th>
<th>unexpected</th>
</tr>
</thead>
<tbody>
<tr>
<td>bimorphemic nouns</td>
<td>The beautiful princess was painted.</td>
<td>The beautifully princess was painted.</td>
</tr>
<tr>
<td>typical nouns</td>
<td>The tasteless soda was marketed.</td>
<td>The tastelessly soda was marketed.</td>
</tr>
<tr>
<td>neutral nouns</td>
<td>The cute infant was dressed.</td>
<td>The cutely infant was dressed.</td>
</tr>
</tbody>
</table>

*Examples of the experimental stimuli of Experiment 3 (40 per condition). See Chapter 6 for a full description.*

If word category violations are detected during early visual processing exclusively on the basis of closed-class morphemes in the input, then only bimorphemic nouns should show an M100 effect of expectedness. Alternatively, if form typicality is sufficient, then an M100 effect should be present for typical nouns as well. Under neither hypothesis should neutral nouns elicit an M100 expectedness effect.

In addition to comparing the averaged M100 responses to each noun type by sentence context, we analyzed dipole waveforms for single-trial data. This allowed us to investigate whether M100 amplitude reflects the extent to which form-feature expectations are met, as would be predicted by the hypothesis laid out in Figure 4.
Paradigm/Task

Participants performed sensicality judgments on sentences such as the ones exemplified in Table 3, presented word-by-word (300ms on/off).

3.3.2 Results

There was an M100 effect of context for both typical nouns and bimorphemic nouns (but not neutral nouns) showing that closed-class morphemes are not crucial for the early detection of word category violations. Furthermore, the multiple-regression analysis on individual trials’ M100 peak amplitude showed that words whose form was less consistent with the predicted word category’s (nouns in the expected context; participles in the unexpected context) generated an increasingly larger M100. This suggests that the M100 generator is sensitive to the extent to which form predictions are satisfied.

Thus, from Experiment 3 we can conclude that the M100 generator is sensitive to form features that are probabilistically associated with word category membership. These results do not strictly speak to whether or not the M100 has access to word category information. It is possible that some form of probability-based category assignment occurs during the M100 response. For example, a particular typical noun could be assigned a .95 likelihood of pertaining to the noun category and .05 probability of being a verb (based on its form features), while a given neutral noun
may be assigned a 50/50 chance of pertaining to either the noun or verb category. This
would render the neutral noun a 50% match in a context where a verb is predicted, as
opposed to 5% for the typical noun. Importantly however, the findings of Experiment
3 are also compatible with a purely top-down form feature matching account, where
the M100 effect reflects a mismatch of a words form features with the predicted form
features. Under such an account, a typical noun does not generate an enhanced M100
where a participle is predicted because it is less likely to pertain to the “participle
category”, but merely because its form features lay far away on the typicality scale
from the predicted form features.

Although these findings cannot distinguish between these two possibilities,
crucially, they allow for the results of Experiments 1 and 2 to be interpreted as a
feature mismatch effect rather than a true word category mismatch: Potentially, the
unexpected bare stems in Experiments 1 and 2, and the unambiguous nouns in
Experiment 2, did not generate an M100 effect because they happened to be category-
neutral on average in terms of their form features (recall that most English nouns and
verbs are ‘typicality neutral,’ so neutral words satisfy the prediction for both
categories). This is confirmed by a reanalysis of a subset of the data of Experiment 2:
When subdividing the unambiguous nouns of Experiment 2 into a typical and a neutral
bin, unexpected typical nouns generated higher M100 amplitudes than expected
typical nouns, and no such effect was observed for the neutral nouns.
3.4 Experiment 4: Syntactic violations only? Or also lexical-semantic violations?

See Chapter 7 for a full description of Experiment 4.

3.4.1 Introduction/Design

The finding that the M100 effect is not restricted to words containing closed-class category marking morphemes raises the question of whether any expectation-based form feature mismatch might generate an M100 effect, including those that are based on lexical-semantic anticipation. Yet, as pointed out in the Introduction, lexical-semantic effects typically occur much later, and localize to temporal and prefrontal regions rather than sensory cortices (see Lau, Phillips, and Poeppel, 2008 for review). I suggested above that this may be due to differences in the relationship between word form and syntactic category as opposed to word form and lexical semantics: A prediction for a participle in a sentence frame like The discovery was solemnly __ can be translated into a prediction for –ed, but how would a prediction for a word denoting a food item after John ate the __ be translated into a form-based prediction? For cases where one particular lexical item is expected (e.g., steak) the prediction does correlate with a specific orthographic sequence steak. Under the hypothesis that linguistic predictions are translated into form-based estimates whenever possible, one would then expect a mismatch effect in these cases. This was tested in Experiment 4.
To investigate whether violations of lexical-semantic predictions may generate early visual effects, we used a very simple picture-noun matching task, where participants were asked to indicate whether a noun phrase accurately described the preceding image. Noun phrases either matched the image or they did not. The function of the pictures was to constrain cloze-probability. For example, a picture of an apple was used to induce a strong prediction for the word *apple* in a subsequent noun phrase initiated by *the ___*. A word like *carrot* then violated this strong prediction. In addition to pictures denoting specific objects, participants were presented with images referring to natural classes of words pertaining to either the animal or food/drink category: A picture of a grocery bag stood for any word describing an edible or drinkable object, and a picture denoting Noah’s Ark functioned as a placeholder for any noun describing an animal. Thus, a grocery bag followed by *the milk* would constitute a match while a noun phrase like *the dog* is a mismatch, and the other way around if these noun phrases followed a picture of Noah’s Ark. The design is presented in Table 4.

If it is true that form-based expectations may be generated in the case of lexical-semantic prediction, but only if a specific lexical item is predicted, then we might expect an M100 effect when comparing matching nouns to non-matching nouns only in those cases where the noun phrase is preceded by a picture denoting a specific object (+PRED in Table 4), and not in cases where any word pertaining to an entire semantic field would constitute a match (-PRED in Table 4).
Table 4 - Examples of experimental stimuli (Experiment 4)

<table>
<thead>
<tr>
<th>+ / - PRED</th>
<th>PRIME</th>
<th>TARGET</th>
<th>+ / - MATCH</th>
<th># of items</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ prediction for specific word (form)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ prediction for specific word (form)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- prediction for specific word (form)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- prediction for specific word (form)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Examples of the experimental stimuli of Experiment 4 (80 per condition). +Predictive conditions were presented twice: once in isolation (Block 1) and once intermixed with -Predictive conditions (Block 2).

The experiment was divided into 2 blocks. Block 1 consisted only of +Predictive trials, while Block 2 contained both +Predictive and –Predictive trials. The +Predictive trials of Block 2 were an exact repetition of the trials that had been presented in Block 1 (see Chapter 7 for the motivations behind this repetition).
Paradigm/Task

Participants engaged in a simple match/mismatch task for the types of stimuli presented in Table 4. The picture was presented for 900ms, followed by the noun phrase (300ms on/off).

3.4.2 Results

As predicted by the hypothesis that the brain generates form-based expectations wherever possible, the results show an increased M100 amplitude for the +Predictive/Mismatch condition compared to the +Predictive/Match condition, albeit only for Block 2. In contrast, there was no visual effect for the –Predictive comparison. This result refutes the hypothesis that syntactic features are always ‘special’ during the earliest stages of analysis. However, it is important to note that this finding does not dismiss all previous findings that lexical-semantic effects typically occur later than morpho-syntactic effects. I only argue that the very early nature of some syntactic effects is likely a reflection of form-based sensory effects that are insensitive to linguistic properties of the stimulus, as now evidenced by the presence of similar effects for lexical-semantic violations.

Syntactic predictions do, however, remain special in that syntactic categories have more straightforward form-correlates than lexical-semantic field do. This is
likely to result in differences between the two types of predictive processing in everyday language comprehension, in stark contrast to the heavily constrained predictive context in the present study. An ecological-validity concern may also explain why the M100 effect was not present during the first part of the experiment: While generating form-based predictions based on lexical-semantic information is possible, the brain might not typically engage in such predictive processing simply because it is not very often necessary or helpful in everyday communicative contexts.

3.5 Summary

The results from Experiment 1-4 demonstrate that sensory cortices are responsive to form-expectations associated with linguistic predictions. These combined findings strongly suggest that when possible, the brain associates linguistic expectations (syntax, lexical-semantics) into form-estimates. Table 5 and Figure 5 summarize the findings discussed above.

The next set of Chapters describes each of these experiments in full.
<table>
<thead>
<tr>
<th>Experiment</th>
<th>Research Question</th>
<th>Design</th>
<th>Result</th>
<th>Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp 1-4</td>
<td>What is the nature of early responses to linguistic violations?</td>
<td></td>
<td></td>
<td>Follow-up question: What is the nature of the representation to which the M100 generator responds? (lexical, syntactic, word form?)</td>
</tr>
<tr>
<td>Exp. 1</td>
<td>Are early visual responses sensitive to phrase-structure violations?</td>
<td>expectedness (confirmed/violated) x word category (preposition/participle/bare stem)</td>
<td>Yes</td>
<td>No lexical access during the M100 response</td>
</tr>
<tr>
<td>Exp. 2</td>
<td>Is an M100 effect obtained for category-unambiguous words with no overt category marking?</td>
<td>expectedness (confirmed/violated) x category index (lexical/closed-class morph/bare stem)</td>
<td>No</td>
<td>No account necessary whereby the M100 generator performs word category analysis based on closed-class morphemes in the input. The M100 seems sensitive to the extent to which form predictions (based on word category predictions) are satisfied.</td>
</tr>
<tr>
<td>Exp. 3</td>
<td>Is the M100 response exclusively sensitive to closed-class category morphology? Or also form typicality?</td>
<td>expectedness (confirmed/violated) x noun type (closed-class morph/bare/neutral)</td>
<td>No Yes</td>
<td>The M100 response is not exclusively responsive to syntactic violations and is thereby likely to reflect pure visual feature matching (i.e., an effect resulting from a mismatch with form expectations made possible via top-down modulation).</td>
</tr>
<tr>
<td>Exp. 4</td>
<td>Is an M100 effect exclusively found for syntactic violations? Or also when there's a mismatch for a semantically highly expected word?</td>
<td>congruence (match/mismatch) x context (+predictive/-predictive)</td>
<td>No Yes</td>
<td></td>
</tr>
</tbody>
</table>
Figure 5 - Summary of results for Experiments 1 - 4

Result graphs for experiments 1 - 4. Dipole analysis of the M100 component in Experiments 1 and 2 revealed an effect for participles (top panel) but not unambiguous nouns, suggesting that the M100 does not support lexical access. Experiment 3 showed that the M100 effect was not restricted to CCMs, opening up the possibility that only form feature analysis takes place during the M100 response: unexpected form typical nouns, but not form neutral nouns generated an M100 effect. Analysis over M100 sensors in Experiment 4 showed that violations of lexical-semantic predictions can also generate an M100 effect, so long as a single lexical item can be predicted.
Chapter 4 - Experiment 1

[Text largely derived from Dikker, Rabagliati, and Pylkkänen, 2009]

4.1 Introduction

Two hypotheses were tested in Experiment 1

(a) Rapid word category identification relies on overt category-marking closed-class morphemes;

(b) Violations of word category predictions affect modality-specific sensory responses.

Participants read sentences word-by-word while their brain activity was monitored with MEG. Each sentence contained a target item whose word category was either expected or unexpected. To test hypothesis (a), the target varied in whether or not category was marked with a closed-class morpheme. Each target item was either a preposition (a free, closed-class morpheme), a regularly inflected participle (a category-marking closed-class morpheme bound to an open-class morpheme), or a bare nominal stem (a free, open-class morpheme). To test hypothesis (b), the current generators of activity at the M100 and the M170 were estimated for each condition with multi-dipole modeling. Root-mean-square (RMS) analyses of sensor data were further used to assess whether the effects obtained in the source waves were also observable in sensor-space.
4.2 Methods

4.2.1 Participants

15 healthy right-handed subjects (7 male) participated. All had normal or corrected-to-normal vision and gave informed consent. All were students or employees at New York University (ages 19 - 39). Two subjects were excluded for not showing typical M100 and M170 components in their individual across conditions grand-average.

4.2.2 Materials

To test whether early effects of unexpected word category are limited to closed-class morphemes, we used three types of manipulations, shown in Table 6.
Table 6 - Examples of experimental stimuli (Experiment 1)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Expected Example sentence</th>
<th>Cloze-P</th>
<th>Unexpected Example sentence</th>
<th>Cloze-P</th>
<th>Nature of the violation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preposition</td>
<td>The boys heard Joe's stories about Africa.</td>
<td>0.441</td>
<td>The boys heard Joe's about stories Africa.</td>
<td>0.000</td>
<td>Presence of an unpredicted free closed-class morpheme.</td>
</tr>
<tr>
<td>Participle</td>
<td>The discovery was reported.</td>
<td>0.297</td>
<td>The discovery was in the reported.</td>
<td>0.002</td>
<td>Presence of an unpredicted bound closed-class morpheme.</td>
</tr>
<tr>
<td>Bare stem</td>
<td>The discovery was in the report.</td>
<td>0.909</td>
<td>The discovery was report.</td>
<td>0.002</td>
<td>Absence of a necessary bound closed-class morpheme.</td>
</tr>
</tbody>
</table>

Examples of the experimental stimuli of Experiment 1. 300 target sentences (60 per condition for Participles and Bare stems, 30 per condition for Prepositions) were intermixed with 240 filler sentences. Cloze-probabilities reflect the proportion of the word category of interest filled in during an off-line sentence completion task (see text): cloze-probabilities for unexpected targets were at or close to zero.

The first comparison used the exact materials of Neville et al.'s original ELAN study, where the target items were unpredicted prepositions, i.e., free closed-class morphemes (Neville et al., 1991). As a second contrast, we tested unexpected participles, which contained the bound closed-class morpheme -ed. This manipulation formed a close English approximation to the German stimuli presented in (2). One obvious difference is that in these German sentences the participial morpheme is a circumfix (ge-STEM-t), rather than a suffix. Note that both the German and the English manipulations involve unexpected rather than ungrammatical targets: in both languages a participial modifier can follow a determiner (so long as this modifier is followed by a noun). In German, the participle should carry a suffixed case marker in order to be grammatical in this context, but under auditory presentation the suffix has
not become available yet at the point where the ELAN is recorded. Thus the ELAN generated by participles of the form ge-STEM-t has to be a response to the prefix rather than the lack of a case-marking suffix. Finally, we tested whether an early response would be dependent on closed-class morphological markers by including a third contrast containing unexpected bare stems which lacked any overt category-marking morphology.

Since we aimed to obtain very early differences between conditions, differences in the lexical material that immediately preceded the target item were of considerable importance. Lau et al. (2006) have pointed out that there may be a serious problem with the baseline of the prepositional manipulation, since the characteristics of the word presented just before the target vary dramatically between conditions (stories vs. John’s). However, Lau et al. found an ELAN response even for prepositions whose baseline was controlled for. Note that our participle and bare stem conditions do not suffer from this possible confound, as they form a perfect cross-over between target (report and reported) and baseline (was and was in the). In other words, as shown in Table 6, the baselines of the expected participles were also the baselines of the unexpected bare stems, and similarly, the baselines of the unexpected participles were identical to the baselines of the expected bare stems. Consequently, results which are not inversed between participles and bare stems are unlikely to result from the intended manipulation.
Example sentences of each condition (6 in total) are presented in Table 1. 300 target sentences (60 per condition for participles and bare stems, 30 per condition for prepositions (replicating Neville et al., 1991) were intermixed with 240 filler sentences. Each subject saw all sentences in a pseudorandom order.

To confirm that our manipulation in fact varied the predictability of word-category, an off-line cloze-probability test was carried out. 60 subjects each read one of four different sets of 140 sentences; each set contained one quarter of the sentences used in the current study (45 sentences), divided such that no subject saw both members of a pair of sentences (e.g., *The discovery was ___ and The discovery was in the ___* ). The remaining sentences in each set consisted of fillers from a separate experiment. Sentences were presented up-to-and-including the pre-target word, and subjects were then asked to write a word, or a short phrase, which they felt completed the sentence. Cloze-probabilities were calculated in terms of the word category of the first word filled in by subjects for each sentence fragment. Note that our definition differs from the typical use of ‘cloze-probability’, which refers to the probability of a specific lexical item. The mean cloze probabilities of the target categories are given in Table 6. The cloze probabilities of the unexpected categories were uniformly at zero, or very close to zero. Thus, across all three manipulations, the probability of the target category was significantly greater (by at least two orders of magnitude) for the Expected than for the Unexpected condition (Prepositions: *t*(58) = 4.1, *p* < 0.001; Participles: *t*(118) = 12, *p*<0.001; Bare Stems: *t*(118) = 62.7, *p*<0.001). These results
confirm for all conditions that unexpected word categories were in fact extremely unlikely continuations of the sentence fragments. For example, in the Unexpected Bare Stems condition, although *was* can be followed by a bare noun, as in *the man was president*, participants did not generate this type of sentences in the cloze probability test.

### 4.2.3 Procedure

During the experiment, subjects lay in a dimly lit, magnetically shielded room. The participants viewed the stimuli through fiberoptic goggles (Avotec, FL). Each trial began with a fixation point in the center of the screen. Participants initiated each trial by pressing a button. The sentences were presented word by word (300ms on, 300ms off), in non-proportional Courier font (font size = 90), and the end of each sentence was indicated with a question mark. At the question mark, participants were instructed to judge whether or not the sentence was well-formed by pressing a button with either the middle or the index finger of their left hand.

Data were collected using a 148-channel whole-head magnetometer (4-D Neuroimaging, Magnes WH 2500), sampling at 678Hz in a band between 0.1 and 200Hz. The entire recording session lasted approximately one hour.
4.2.4 Analysis

Prior to averaging, the MEG data were cleaned of artifacts and trials on which participants provided an incorrect well-formedness judgment. On average, this resulted in the exclusion of 18% of the data per subject (SD = 7%). Data were averaged by stimulus category over a 1300ms epoch with a 300ms pre-stimulus interval, time-locked to the appearance of the target word. Prior to analysis, the recordings were high and low-pass filtered at 1 and 40 Hz respectively.

Dipole Modeling

In order to obtain a maximally complete characterization of the neural sources activated by our stimuli during the first 200ms, the current generators of all prominent response components at 0-200ms were estimated with a multiple-source model using BESA (Brain Electrical Source Analysis 5.1). Subjects typically showed two major peaks in this time-window, associated with the M100 and M170 field patterns (see Fig. 1 for grandaveraged sensor waveforms from all subjects). Thus we created two multiple source solutions, aimed at characterizing all activity present at the M100 and M170 components. BESA minimum norm estimates (a distributed source model) were used to guide the hypothesis about how many sources were active at each component and data from all sensors were then used to estimate the discrete locations of these generators. In other words, we did not use sensors of interest but rather aimed to
obtain a global characterization all early activity. Thus although our hypothesis pertained to the early visual activities, our method was equally suited for identifying any other relatively focal early sources, such as potential left anterior generators.

Dipole modeling is often performed using each individual’s grandaveraged data across conditions (e.g., Pylkkänen et al., 2006), to ascertain that the models are based on data with a maximally high signal-to-noise ratio. This procedure, however, was less than ideal for our data: Although it should give an accurate characterization of sensory activity - and thus allow us to address our main hypothesis - it risks missing non-sensory ELAN sources that may not be active across all conditions. Therefore we aimed to model activity using data from the individual conditions. This was successful for the M100 component, but less so for the M170, whose bilateral sources can be difficult to distinguish with this type of multi-dipole modeling (cf., Zweig and Pylkkänen, 2009). To enhance signal-to-noise ratio, we instead modeled the activity at the M170 using the combined data from the Expected and Unexpected conditions for each word category (resulting in three M170 models per subject: one for the Prepositions grandaverage, one for the Participles grandaverage, and one for the Bare Stems grandaverage). This model was then applied to the Expected and Unexpected conditions.

Only models that were consistent with both the magnetic field maps and the minimum norm estimates were accepted for analysis. This resulted in the inclusion of 11 subjects in the analysis of the M100 time-window. In the M170 time-window, 9
subjects’ dipole models were accepted for the participles and the bare stems, but only 6 subjects’ for the prepositions. The average source locations of all M100 and M170 dipoles are shown in Figures 7 and 8 respectively.

Our multi-dipole modeling did not yield any consistent source clusters apart from the M100 and the M170 generators. In other words, although other dipoles were present in many subjects’ models, their locations did not exhibit any observable generalization across subjects. Thus statistical analysis was only possible on the M100 and M170 dipoles. M100 and M170 peaks were first identified for each condition using across-subjects grandaveraged source waveforms. After this, an interval of interest, 15ms around the peak, was defined for each condition. Dipole strength was calculated and averaged across this interval per subject and condition. These values were then entered into a 2 (Expectedness: Expected vs. Unexpected) by 3 (Word Category: Preposition vs. Participle vs. Bare Stem) within-subjects ANOVA. Separate ANOVA’s were performed for the M100 and M170 components. Planned comparisons tested for the effects of expectedness for each target word category.

RMS Analysis

A root-mean-square analysis of sensor data was used as a second analysis to further assess the robustness of the findings from the multi-dipole modeling analysis. This
also allowed us to test whether left-anterior effects might be obtained in sensor data. This analysis was performed over 5 equally sized regions of interest (ROI’s) (see Figures 9 and 10), each containing approximately 18 sensors. To test for visual M100 effects, we examined two posterior ROI’s, which captured the field pattern of the M100 component (the right-hemisphere posterior ROI captured the outgoing magnetic-field produced by the M100 generator, the left-hemisphere posterior ROI captured the re-entering field). To investigate possible activity in left temporal and frontal areas, the remaining left-hemisphere sensors were also examined. They were divided into three ROI’s in order to differentiate between more posterior and anterior sources. The regions and sensors within them were held constant across subjects to ensure a consistent analysis.

RMS waveform analysis followed the same procedure used in the source wave analysis, described above. To analyze activity in each ROI, the averaged root-mean-square value of the sensor activity in the ROI was first calculated for each subject in each condition. For each condition, the grand-average RMS wave of all subjects was then examined to identify peak activity. This revealed a consistent and prominent peak in the M100 time-window (100-150ms). Our RMS analysis focused on this activity. Activity was averaged across a 15ms window around this peak, per condition, per subject and per ROI, and then entered into a 2 (Expectedness: Expected vs. Unexpected) by 3 (Word Category: Preposition vs. Participle vs. Bare Stem) within-subjects ANOVA.
For the most frontal ROI, no clear peak activity in the time-window of interest was identified. In this case, we also averaged each subject’s RMS data, for each condition, over the window between 100ms and 150ms. This was then also entered into a 2 (Expectedness: Expected vs. Unexpected) by 3 (Word Category: Preposition vs. Participle vs. Bare Stem) within-subjects ANOVA, in the same way as for the other measure.

4.3 Results

4.3.1 Behavioral Data

Overall accuracy was very high; 96% of sentences were correctly classified in the well-formedness judgment task.

4.3.2 MEG Data

Figure 6 visualizes the across participants (n=13) grandaveraged MEG signals for each of the six conditions. The magnetic field patterns associated with sensor activity during the first peak (indicated by the dashed line) showed a typical M100 distribution in all conditions. Visual inspection of this peak revealed that its amplitude increased for unexpected items containing a closed-class morpheme (participle and prepositions), as compared to the expected items. However, there appeared to be no
change in M100 amplitude for unexpected items not containing a closed-class item (the bare stems). The second peak, around 170ms, represents the M170 activity. Note that this peak was less consistent across conditions for these 13 subjects. The preposition conditions in particular did not show a clear peak.

Figure 6 - Grandaveraged waveforms and field distributions of the M100 (Experiment 1)

Grandaveraged waveforms and field distributions (n=13) for each condition. All sensors are shown. In the waveforms, the cursor is pointed at the largest peak within 100-150ms. The gradient maps on the left and right depict the magnetic field patterns associated with these peaks (10fT/line). In each case, the field pattern is a canonical M100 pattern. Both the waveforms and the field maps suggest an effect of expectedness at the visual M100 for unexpected prepositions and participles, but not for the bare stems.
4.3.3 Dipole Results

4.3.3.1 The M100 time-window

Consistency in the multi-dipole models

Dipole solutions for the eleven subjects included in the analysis contained a typical posterior M100 dipole in all conditions (Tarkiainen et al., 1999). Each subject also required other dipoles to be included in their solutions in order for the M100 dipole to localize accurately. There was no consistent pattern in the location/direction of these additional dipoles, and no reliable difference between conditions with regard to the number of additional dipoles. Average goodness of fit for the multi-dipole models was very high, greater than 90% in every condition (overall mean 92.5%, SD = 6.0), and did not reliably vary between conditions.

To test whether conditions differed in M100 dipole location or orientation, we ran 2 (Expectedness: Expected vs. Unexpected) by 3 (Word Category: Preposition vs. Participle vs. Bare Stem) within-subjects ANOVA’s for the x, y and z location and orientation measures (on a Cartesian coordinate system). These analyses revealed no reliable main effects or interactions, except for a reliable interaction between Word Category and Expectedness on location along the x-axis ($F(2,20) = 3.81, p = 0.04$). However, these differences were not significant when the effect of Expectedness was analyzed separately per word category.
Effects of Expectedness at the M100

The averaged waveforms of M100 dipole strength over time are shown in Figure 7.

**Figure 7 – M100 dipole waveforms and peak amplitudes (Experiment 1)**

Grandaveraged waveforms for the M100 dipole sources per comparison (n=11) and mean amplitudes in nAm for the 15ms intervals centered around the M100 peaks (the time-window between the dotted lines in the waveform graphs). Mean dipole locations and orientations (blue = expected / red = unexpected) as well as the dipoles from the individual participants (grey) are plotted per word category. Results reveal effects of expectedness on M100 amplitude, but only for the Prepositions and Participles (* = p < .05).
M100 peak amplitudes were entered into a 2 (Expectedness: Expected vs. Unexpected) by 3 (Word Category: Preposition vs. Participle vs. Bare Stem) within-subjects ANOVA. This revealed no main effect of Word Category ($F(2,20) = 1.62, p = 0.22$), a marginal main effect of Expectedness ($F(1,10) = 4.65, p = 0.056$), and a reliable interaction between Word Category and Expectedness ($F(2,20) = 4.22, p = 0.030$).

Planned comparisons revealed that, for the Preposition and Participle conditions, Unexpected items generated a reliable increase in M100 amplitude (Prepositions: $t(10) = 2.45, p = 0.034$. Particiles: $t(10) = 2.46, p = 0.033$). However, there was no reliable effect for the Bare Stem conditions ($t(10) = 1.03, p = 0.33$). Thus these results support both the sensory hypothesis as well as the hypothesis that early effects of category violations are limited to closed-class morphemes.

### 4.3.3.2 The M170 time-window

**Consistency in the multi-dipole models**

For the bare stems and the participles, the dipole-models of 9 subjects included two (bilateral) dipoles whose location, orientation, and latency was typical for the M170 component (Tarkiainen et al., 1999). For the prepositions, this was the case for only 6 subjects.

Eight subjects’ models required additional, non-M170-like, dipoles to be included in the solutions (2 dipoles on average, SD = 1.2, maximum: 4 additional
dipoles for 2 subjects). Again, there was no consistent pattern in the location/direction of these dipoles. Average goodness of fit for the multi-dipole models was high, greater than 90% for the participle and bare stem conditions, and greater than 85% for the preposition conditions (overall mean 92.3 %, SD=7.1), and did not vary reliably between conditions.

Because activity was modeled on the basis of the combined expected and unexpected data for each word category, and because the number of subjects included in the analysis varied across word categories, we did not run statistical analyses on the locations and orientations of the M170 dipoles. However, visual inspection of the dipoles in Figure 8 suggests a high consistency in both location and orientation.

Effects of Expectedness at the M170

The averaged waveforms of M170 dipole strength over time are shown in Figure 8. Because the M170 is generated by a pair of bilateral sources, average source-waveform activity centered around the M170 peaks was entered into a 2 (Hemisphere: Left vs. Right) by 2 (Expectedness: Expected vs. Unexpected) within-subjects ANOVA for the prepositions, and a 2 (Hemisphere: Left vs. Right) by 2 (Expectedness: Expected vs. Unexpected) by 2 (Word Category: Participle vs. Bare Stem) within-subjects ANOVA for the participles and bare stems.
Figure 8 - M170 dipole waveforms and peak amplitudes (Experiment 1)

Grandaveraged waveforms for the left and right M170 dipole sources per condition and mean amplitudes in nAm for the 15ms intervals centered around the left and right M170 peaks (the time-window between the dotted lines in the waveform graphs). Mean dipole locations and orientations (black) as well as the dipoles from the individual participants (grey) are plotted in the center for each word category. Results reveal no effects of expectedness on M170 amplitude.
As can be seen in Figure 8, M170 amplitudes appeared larger for unexpected than for expected prepositions, but this effect was not reliable: the prepositions showed no main effects and no interactions. The participles and bare stems also showed no main effects, but they did exhibit a reliable Hemisphere by Word Category interaction ($F(1,8) = 6.37, p = 0.036$), as well as a reliable Expectedness by Word Category interaction ($F(1,8) = 14.68, p = 0.005$).

To explore these interactions, we ran follow-up 2 (Hemisphere: Left vs. Right) by 2 (Expectedness: Expected vs. Unexpected) within-subjects ANOVAs separated by word category. This showed no effects for the Bare Stem comparisons. For the participles, there was a marginal main effect of Hemisphere ($F(1,8) = 5.12, p = 0.054$), with the Left Hemisphere M170 dipole showing more activity than the Right Hemisphere M170 dipole. There was no main effect of Expectedness, nor did Expectedness interact with Hemisphere. A 2 (Word Category: Participle vs. Bare Stem) by 2 (Expectedness: Expected vs. Unexpected) within-subjects ANOVA by hemisphere showed no effects for the Right Hemisphere M170 source. For the Left Hemisphere Source, there were no main effects, but there was an interaction between Word Category and Expectedness ($F(1,8) = 6.60, p = 0.033$), suggesting that the Unexpected Participles yielded a higher M170 amplitude than the Expected Participles, while the reverse was true for the Bare Stem conditions. However, post-hoc t-tests revealed no reliable difference by Expectedness for either the Participles ($t(8) = 1.06, p = 0.322$), or the Bare Stems ($t(8) = 1.66, p = 0.136$).
4.3.3.3 Anterior Sources?

In addition to the M100 and M170 generators, we would have expected to capture any consistent left anterior sources clusters, at least in the unexpected conditions, if these stimuli were associated with a left-anterior effect. However, there was no consistent grouping of dipoles in anterior regions of the left-hemisphere. This could indicate that there were no left-anterior discrete sources. However, it may also be because the sensory components dominate our time-window of interest. Especially the visual M100 is usually reflected by a very strong magnetic field pattern covering most of the sensors. Since we did not perform dipole analyses based on sensors-of-interest but rather let the dipole locations and orientations vary freely, a weaker left-anterior field may have been masked by the M100. To further test for left anterior effects, and to confirm the posterior effects, we carried out an RMS analysis of posterior and left anterior ROI’s.

4.3.4 RMS Results

4.3.4.1 Posterior Octants

To capture the ingoing and outgoing field maxima of the visual M100, we first collapsed the RMSs of the two most posterior regions of interest (Figure 9A). Surprisingly, there was no significant main effect of Expectedness or Word Category, nor a reliable interaction between the two.
Given that the M100 is bilaterally generated, we then tested for potential effects of laterality by analyzing the two posterior ROIs separately (Figures 9B and 9C). This analysis did reveal a significant main effect of Expectedness in the left posterior ROI ($F(1,12) = 5.19, p = 0.042$), showing increased amplitudes for the expected items. No significant main effect of Word Category ($F(2,24) = 1.46, p = 0.25$) and no significant interaction between Expectedness and Word Category ($F(2,24) = 0.65, p = 0.53$) were obtained. However, planned comparisons revealed that the main effect of Expectedness on peak amplitude was driven by those conditions containing overt closed-class morphemes. In other words, reliable increases in amplitude were found for the Preposition condition ($t(12) = 2.31, p = 0.04$) and the Participle condition ($t(12) = 2.18, p = 0.049$), but not for the Bare Stem condition ($t(12) = 0.47, p = 0.65$).

In contrast, the right posterior ROI (Figure 9C) displayed no effect for any of the manipulations, nor an interaction. Planned comparisons also revealed no reliable effects of Expectedness for any word category. However, the Participles did show a marginal effect of Expectedness, but this was due to an increase in activation in the expected condition ($t(12) = -2.09, p = 0.059$).

In summary, analyses of the posterior ROIs revealed a left-lateralized effect of Expectedness at the visual M100 peak. In the planned comparisons, this effect was limited to the preposition and participle contrasts, i.e., to targets containing either a free or a bound overt closed-class category-marking morpheme. This pattern of
findings is consistent with the dipole-analysis – the mean dipole locations across conditions were medial, but slightly left-lateralized.

**Figure 9 - sensor waveforms and peak amplitudes in posterior octants (Experiment 1)**

*Left to right:* Grandaveraged RMSs per condition and mean amplitudes in fT for the 15ms intervals of interest (indicated by the dotted lines in the RMS waveform graphs).

*The three regions-of interest (ROIs), top to bottom:* the two posterior octants combined (A), the left posterior octant (B), and the right posterior octant (C). ROIs are highlighted in sensor space. Results reveal reliably enhanced activity for the unexpected conditions in the left posterior ROI only (* = p < .05; n = 13).

**4.3.4.1 Anterior Octants**

To assess the possible contribution of additional left anterior generators to the effects of expectedness, we analyzed the remaining three left lateralized regions of interest in the same manner as above.
Figure 10 - Sensor waveforms and peak amplitudes in LA octants (Experiment 1)

Left to right: Grandaveraged RMSs per condition and mean amplitudes in fT for the 15ms intervals of interest (indicated by the dotted lines in the RMS waveform graphs).

Three regions-of-interest (ROIs), top to bottom: the left medial-posterior (A), left medial-anterior (B), and left-anterior octant (C). ROIs are highlighted in sensor space.
In the medial-posterior ROI (Figure 10A), clearly overlapping with the M100 field distribution, peak amplitude in the time-window of interest showed a marginal main effect of Expectedness ($F(1,12) = 3.89, p = 0.072$), but no main effect of Word Category or reliable interaction between the two. Planned comparisons revealed a significant effect of Expectedness for the Preposition condition ($t(12) = 2.57, p = 0.025$), but no other effects. In the medial-anterior ROI (Figure 10B), capturing only the very edges of the M100 distribution, peak amplitude in the 100-150 ms time-window no longer showed any significant main effects or interactions. Again, planned comparisons revealed an effect of Expectedness for the Preposition condition ($t(12) = 2.32, p = 0.039$) but no other effects.

The most anterior ROI (Figure 10C) was the most crucial for any potential left anterior effects: while for the other ROIs examined so far, the modulation found for the Preposition conditions could be due to spill-over from a posterior M100 effect, a posterior generator alone would be unlikely to explain an effect in this ROI. Thus an effect in this ROI would strongly suggest an additional left-anterior generator. However, analysis of this ROI’s peaks in the 100-150 time-window showed no reliable main effects of Expectedness or Word Category, and no reliable interaction between the two. In addition, planned comparisons revealed no significant differences between the conditions. The analysis of this ROI was complicated by the low amplitude of the responses, which impeded detection of peaks. To ensure our result was not an artifact of this, we also performed an identical ANOVA for the mean
amplitude of the 100-150ms interval, as explained in Section 4.2.4. This revealed a significant main effect of Word Category ($F(2,24) = 6.73, p = 0.005$), Particiles showing a higher amplitude than the other word categories overall, but no other effects.

### 4.3.5 Summary of Results

The peak amplitude of the visual M100, a neuromagnetic sensory response to visual stimulation, was modulated by the expectedness of target category, but only when the target was overtly marked for word category by a closed-class morpheme. This effect of expectedness was found in both the M100 dipole and RMS analyses. The peak amplitude of the subsequent M170 component demonstrated no effects of Expectedness in any of the conditions. There was a marginal effect of Hemisphere for the Particiles, showing increased activity for the Left as compared to the Right M170 dipole.

Neither the dipole nor the sensor analysis showed evidence for anterior generators.
4.4 **Discussion**

Experiment 1 investigated two hypotheses about the generators and generating conditions of early neural responses to unexpected word category violations in sentential contexts. Using magnetoencephalography, we tested whether early syntactic responses depend on the presence of category-marking closed-class morphemes in unexpected target words and whether these early responses are at least partly generated by enhanced activity in sensory cortices. Consistent with both hypotheses, we found the peak amplitude of the visual M100 response to be modulated by expectedness but only when the target word contained overt category-marking functional morphology.

Multi-dipole modeling suggested a posterior generator for this effect, and did not reveal further left-anterior sources in any conditions. RMS analyses of regions of interest also revealed a posterior effect, although this effect was left lateralized. No effect was found in the most anterior region of interest. Further, no reliable modulation of the M100 was found for the morphologically unmarked bare stems in any ROI.

The visual M170 response, a slightly later component that has been associated with the Visual Word Form Area, did not show any effects of expectedness for any of the word categories. Thus in this experiment, we found no evidence suggesting that the M170 participates in the post-M100 computation of category violations. However, due to the relative difficulty in modeling the M170, this conclusion must remain
tentative. For example, in the preposition manipulation, there was a numerical M170 difference between the expected and unexpected conditions, but this difference was not reliable, potentially because of the small number of subjects for which a clear bilateral M170 source could be modeled.

Two factors in particular partially confound the conclusion that the null result for the bare stems was due to the lack of closed-class morphology. First, unlike the participles and prepositions, the word category of the bare stems was ambiguous between noun and verb. Second, the expectations violated by the unexpected participles and the unexpected prepositions were considerably stronger than the expectations violated by the unexpected bare stems (see cloze-probabilities in Table 7). Experiment 2 tests whether either of these two factors was responsible for our null results for the bare stems.
Chapter 5 - Experiment 2

[Text largely derived from Dikker, Rabagliati, and Pylkkänen, 2009]

5.1 Introduction

Experiment 2 was designed to assess whether the bare stems of Experiment 1 failed to show an M100 modulation because of their category ambiguity or because of the lesser contrast in prediction strength between the expected and unexpected conditions.

As already observed above, in Experiment 1 all bare stems were ambiguous for word category (nominal/verbal). It is possible that this ambiguity, rather than the lack of closed-class morphology, resulted in the absence of an M100 effect. To test whether this may have confounded the results of Experiment 1, Experiment 2 included unambiguously nominal stems like ‘tree’, in addition to the ambiguous stems tested in Experiment 1 (e.g., ‘report’):

(3)   a. The owl was in the TREE.
    b. * The owl was TREE.

In addition, the unexpected bare stem condition in Experiment 1 violated a weaker prediction for a specific word category than the unexpected participle condition.
Recall that Lau et al. (2006) found that the ELAN is sensitive to the strength of a violated prediction. The M100 modulation may thus be dependent upon violating a particularly strong prediction, and the bare stem condition in Experiment 1 may not have met this criterion. For the sentence fragment preceding the target word in the unexpected bare stem condition (e.g., *The discovery was ___*), the cloze-probability of a participle was only 29% and there was no strong prediction for any other category either. Thus the unexpected bare stems contrasted with the unexpected participle condition (e.g., *The discovery was in the *reported*), where the cloze-probability of a noun was 91%. This discrepancy raises the question whether the relatively low prediction strength in the unexpected bare stem condition may have caused the lack of M100 modulation.

To investigate this, we created a very strong prediction for a participle by inserting an adverb into the stimuli of Experiment 1 (e.g., *The discovery was solemnly report*). As discussed in Experiment 1, it is important to control for possible effects that may be due to the baseline of a target word rather than the intended manipulation. In Experiment 1, we solved this problem by making sure that in each stimulus the baseline for an expected word (e.g., ___ was reported) was also the baseline for an unexpected item in another condition (e.g., ___ was report) and vice versa (e.g., ___ in the report / ___ in the reported). To ensure that a similar baseline cross-over held for the conditions in Experiment 2, we also inserted an adjective in the expected bare stem/noun conditions (e.g., *The discovery was in the solemn ___*; see section 3.2.2 for
details). In addition to these conditions, all other conditions of Experiment 1, except the prepositions, were also included in Experiment 2.

In summary, in Experiment 2 we added two additional factors to the design of Experiment 1, namely category ambiguity and prediction strength, in order to investigate whether M100 effects are limited to close-class morphemes or modulated by these other factors.

5.2 Methods

5.2.1 Participants

12 healthy right-handed subjects (3 male) participated. All had normal or corrected-to-normal vision and gave informed consent. All were students or employees at New York University (ages 19-42).

5.2.2 Materials

Our materials consisted of three types of manipulations. First we tested unexpected participles containing the bound closed-class morpheme –ed, exactly as in Experiment 1, with the exception of a few items that were replaced because they did not lend themselves to the addition of an adjective or adverb (see below). Second, we included
the bare stem manipulation as in Experiment 1, but with the added participles. Again, as in Experiment 1, the only difference between the participle conditions and the bare stem conditions was the presence of the –ed morpheme. Finally, we included sentences that contained unambiguous nouns (items defined as nouns by the COBUILD corpus, and which COBUILD listed as either not having a verbal form, or whose verbal form’s COBUILD frequency was below 750 and rarer by at least two orders of magnitude than the noun form). These unambiguous nouns were open class, but were not productively used as any other word category. This condition allowed us to test whether the presence of closed-class morphology is indeed a prerequisite for the M100 effect, or whether this effect is also obtained for unexpected open-class morphemes that are category unambiguous.

To investigate whether the absence of an M100 effect for the bare stems in Experiment 1 was caused by a lack of a strong category prediction, prediction strength was varied across word category, including in the participle manipulation, by inserting either an adjective or a manner adverb. As pointed out above, the intention of inserting this additional material was to create a very strong prediction for a participle after sentence fragments such as the discovery was solemnly. The full design (6 conditions total) is shown in Table 7. 60 sentences were presented in each condition (720 sentences overall).
Table 7 - Examples of experimental stimuli (Experiment 2)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Word Category</th>
<th>Category Ambiguity</th>
<th>Prediction Level</th>
<th>Example sentence Expected</th>
<th>Cloze-P</th>
<th>Example sentence Unexpected</th>
<th>Cloze-P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participle</td>
<td>Unambiguous</td>
<td>Weak</td>
<td>The discovery was reported.</td>
<td>0.27</td>
<td>The discovery was in the reported</td>
<td></td>
<td>Part: 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Strong</td>
<td>The discovery was solemnly reported.</td>
<td>0.79</td>
<td>The discovery was in the solemn reported</td>
<td></td>
<td>Part: 0</td>
</tr>
<tr>
<td>Bare stem</td>
<td>Ambiguous</td>
<td>Weak</td>
<td>The discovery was in the report.</td>
<td>0.88</td>
<td>The discovery was report.</td>
<td></td>
<td>Noun: 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Strong</td>
<td>The discovery was in the solemn report.</td>
<td>0.90</td>
<td>The discovery was solemnly report.</td>
<td></td>
<td>Noun: 0</td>
</tr>
<tr>
<td>Noun</td>
<td>Unambiguous</td>
<td>Weak</td>
<td>The owl was in the tree.</td>
<td>0.79</td>
<td>The owl was tree.</td>
<td></td>
<td>Noun: 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Strong</td>
<td>The owl was in the high tree.</td>
<td>0.89</td>
<td>The owl was highly tree.</td>
<td></td>
<td>Noun: 0</td>
</tr>
</tbody>
</table>

Examples of the experimental stimuli of Experiment 2 (60 per condition). Cloze-probabilities reflect the proportion of the word category of interest filled in during an off-line sentence completion task (see text). For the unexpected conditions, cloze-probabilities are also given for the violated word category expectation (i.e., participles for the unexpected bare stems and the unexpected unambiguous nouns, and nouns for the unexpected participles).

Cloze-probability

To confirm that the different conditions manipulated predictions for the target word-category, an off-line cloze-probability test was carried out, as in Experiment 1 (see Section 2.2.2). Cloze-probabilities were calculated in terms of the word category of the first word filled in by subjects for each sentence fragment. Results are given in Table 2.

As in Experiment 1, the probability of the word category of the target word was significantly greater for the Expected than for the Unexpected conditions across all
three manipulations. Importantly, for each Word Category, there was a significant effect of Prediction Strength on cloze-probability: A participle was significantly more often filled in after an adverb than after a auxiliary ($t(118) = 14.13, p < 0.001$), and a noun was more often filled in after an adjective than after a determiner (Bare Stems: $t(118) = 3.17, p = 0.002$; Nouns $t(118) = 5.16, p < 0.001$). Notice, however, that the percentage point difference between the likelihood of a participle after an adverb vs. after an auxiliary was much greater (27% vs. 79%) than the difference between the likelihood of a bare stem after an adjective vs. after a determiner (86% vs. 90%); this interaction between our prediction strength manipulation and condition was reliable ($F(1,236) = 150.3, p < 0.001$). Our manipulation therefore succeeded in increasing the strength of the violated prediction in the Unexpected Bare Stem and the Unambiguous Noun manipulations.

5.2.3 Procedure

During the experiment, subjects sat in a dimly lit, magnetically sealed chamber. The participants viewed the stimuli on a screen approximately 17 inches from their head. Each trial, initiated by participants via button press, began with a fixation point in the center of the screen. The sentences were presented word by word (300ms on, 300ms off), in non-proportional Courier font (font size = 90), and the end of each sentence was indicated with a question mark. At the question mark, participants were instructed
to judge whether the sentence was well-formed or not by pressing a button with either the middle or the index finger of their left hand. The entire recording session lasted approximately 75 minutes.

Data were collected using a whole-head 275-channel gradiometer (CTF, Vancouver Canada) system sampling at a 600Hz in a band between 0.1 and 200Hz. Since Experiments 1 and 2 used different MEG machines, our study also serves to address to what extent the results of Experiment 1 replicate with a different machine.

5.2.4 Analysis

Prior to averaging, the MEG data were cleaned of artifacts. On average, this resulted in the exclusion of less than 10 % of the data per subject. Data were averaged by stimulus category over a 900ms epoch with a 300ms pre-stimulus interval, time-locked to the appearance of the target word. Prior to analysis, the recordings were high and low-pass filtered at 1 and 40 Hz respectively.

M100 activity was modeled for each condition and each subject separately following the procedure outlined for Experiment 1, in section 4.2.4.
5.3 Results

5.3.1 Behavioral Data

Overall accuracy in the grammaticality judgment task was very high, averaging at 94%.

5.3.2 M100 Dipole Results

M100 multi-dipole models

For 5 of the 12 subjects, the M100 activity could be modeled by a single dipole in all conditions. All other subjects required at least one additional dipole. The mean number of dipoles for all subjects was 1.4; there were no reliable differences in the number of dipoles between conditions. As in Experiment 1, there was also no consistent pattern with respect to Expectedness or Word Category for either the number or location of the additional non-M100 dipoles. The mean goodness of fit for the dipole solutions was 89.5%. (SD = 5.9%). Figure 6 shows the average waveform and average dipole for each condition (blue = expected / red = unexpected), as well as the individually modeled dipoles (grey).

To investigate possible consistent variation in dipole location and orientation across conditions, the x, y, and z locations and orientations of the individual M100
dipoles per condition were entered into a series of 2 (Expectedness: Expected vs. Unexpected) by 2 (Prediction Strength: Weak vs. Strong) by 3 (Word Category: Participle vs. Bare Stem vs. Unambiguous Noun) within-subjects ANOVAs.

For dipole location, an interaction was observed between Word Category and Expectedness along the anterior-posterior axis (y-location) ($F(2,22) = 5.278, p = 0.021$). For the Participles the generators of the M100 for Unexpected Participles were significantly more anterior (-.676) than for the Expected Participles (-.725) overall ($F(1,11) = 11.328, p = 0.006$). In the Unambiguous Noun comparison, in contrast, the M100 dipoles for Unexpected conditions were significantly more posterior (-.678mm) than for Expected conditions (-.652mm) overall ($F(1,11) = 5.737, p = 0.036$).

However, as can be seen in Figure 11, these location differences were extremely slight, smaller than the resolution of MEG (0.04mm for the Participles and 0.02mm for the Unambiguous Nouns), and do not warrant any strong conclusions.

Dipole orientation vectors also demonstrated a very slight (less than .04) but reliable main effect of Expectedness on the y-coordinate ($F(1,11) = 5.527, p = 0.038$; Expected: -.223; Unexpected: -.186). No further effects on dipole orientation or location were found.
Figure 11 - M100 dipole waveforms and peak amplitudes (Experiment 2)
Grandaveraged waveforms for the M100 dipole sources per comparison (n=12) and mean amplitudes in nAm for the 15ms intervals centered around the M100 peaks (time-window between dotted lines in graphs). Mean dipole locations and orientations (blue = expected / red = unexpected) as well as the dipoles from the individual participants per condition (grey) are plotted.

Effects of Expectedness at the M100
Following the same procedure as in Experiment 1, amplitudes of the M100 component over a 15ms interval (9 time samples) centered around the component’s peak (indicated with dotted lines in Figure 6) for each condition were entered into several within-subjects ANOVAs. A 2 (Expectedness: Expected vs. Unexpected) by 2
(Prediction Strength: Weak vs. Strong) by 3 (Word Category: Participle vs. Bare Stem vs. Unambiguous Noun) within-subjects ANOVA revealed a significant main effect of Expectedness ($F(1,11) = 4.93, p = 0.048$) as well as a reliable interaction between Expectedness and Word Category ($F(2,22) = 5.08, p = 0.028$). Follow-up 2 (Expectedness: Expected vs. Unexpected) by 2 (Prediction Strength: Weak vs. Strong) ANOVAs separated by Word Category were carried out to clarify this interaction.

First, the Participle comparisons revealed a main effect of Expectedness ($F(1,11) = 8.82, p = 0.013$): Unexpected Participles yielded higher M100 amplitude than the Expected Participles overall. There was no main effect of Prediction Strength ($F(1,11) = 2.13, p = 0.17$), nor was there an interaction between Expectedness and Prediction Strength ($F(1,11) = 0.51, p = 0.49$). In other words, the Expectedness effect was not modulated by Prediction Strength.

In contrast to the Participle conditions, the Bare Stems manipulations showed no main effect of Expectedness ($F(1,11) = 0.002, p = 0.97$), but there was a marginal main effect of Prediction Strength ($F(1,11) = 4.29, p = 0.063$), which did not interact with Expectedness ($F(1,11) = 0.002, p = 0.96$). In other words, participants showed higher amplitudes for contexts that set up strong predictions, irrespective of whether this prediction was violated or satisfied. The Unambiguous Nouns showed a similar pattern as the Bare Stems (Expectedness: $F(1,11) = 0.22, p = 0.65$; Prediction Strength: $F(1,11) = 13.66, p = 0.004$; Expectedness x Prediction Strength: $F(1,11) =
0.099, \( p = 0.76 \)). Thus, as in Experiment 1, Expectedness only modulated the M100 generated by items containing closed class morphemes.

Finally, the overall ANOVA including all conditions and word categories revealed a main effect of Prediction Strength (\( F(1,11) = 14.72, p = 0.003 \)), which did not interact with Expectedness (\( F(1,11) = 0.216, p = 0.651 \)), or Word Category (\( F(2,22) = 0.65, p = 0.889 \)): Participants displayed increased M100 amplitudes to targets appearing in strongly predictive contexts, irrespective of word category or whether the expectation was violated. That targets satisfying a strong prediction and targets violating a strong prediction should both be associated with the same M100 amplitude increase is quite surprising and certainly does not follow from any of the hypotheses we tested in this work. This effect was robust though and thus merits further investigation.

### 5.4 Discussion

Experiment 1 demonstrated that the amplitude of the visual M100 is modulated by the presence of an unexpected word category, but only when the target item contains a closed-class category marking morpheme. To assess whether the closed/open class distinction was in fact the critical factor, Experiment 2 tested two alternative explanations of this effect, one based on category ambiguity and the other on prediction strength. Specifically, we tested whether we would obtain an M100 effect
for category unambiguous nouns which lack closed-class category-marking morphology and violate a strong prediction for a participle.

Category disambiguation and the increase in prediction strength did not result in an M100 effect for the category unambiguous nouns. Bare stems also failed to show an effect, as in Experiment 1. The only word category showing an M100 effect of expectedness was the Particles, i.e., the items containing closed-class category marking morphology. This appears to support the hypothesis that closed-class morphology is indeed the crucial factor. However, this conclusion is tentative, as will be shown in the next Chapter.

Surprisingly, we obtained a main effect of prediction strength – both strongly expected and unexpected items showed increased M100 amplitudes as compared to the weak prediction conditions. In other words, overall higher M100 amplitudes were found for participle, bare stem and noun targets after adverbs and adjectives than after determiners and the copula. Thus prediction strength does affect the M100, although perhaps in a somewhat complicated way.
Chapter 6 - Experiment 3

[Text largely derived from Dikker, Rabagliati, Farmer and Pylkkänen, 2010]

6.1 Introduction

Experiments 1 and 2 revealed an M100 effect only when the category of the unexpected item was saliently marked by a closed-class morpheme, making these obvious candidates for the relevant category cues that the occipital cortex responds to. However, as laid out in Figure 3, under the hypothesis that closed-class morphemes are necessary, the pattern of data reported for Experiments 1 and 2 force us to conclude that the M100 generator supports not only morpheme form analysis, but also word category access, based on the presence of closed-class morphemes in the input.

Alternatively, however, the relevant category cues could be sets of probabilistic form features that are indicative of a particular syntactic category. In that case, an M100 effect of unexpectedness would be obtained even for words that lack a closed-class morpheme, as long as their form is characteristic of their syntactic category. The M100 findings on closed-class morphemes reported in the chapters above could easily be explained by this hypothesis because a word with a category-marking morpheme is very likely to look typical of its category.

This form-typicality hypothesis derives from research demonstrating that the words of a given syntactic category share systematic, probabilistic, form-based
regularities, and that these regularities have consequences for on-line syntactic processing (Farmer, Christiansen, and Monaghan, 2006; Kelly, 1992; Monaghan, Christiansen, and Chater, 2007). In one recent study, Farmer et al. (2006) demonstrated via a corpus analysis that English nouns and verbs form clusters in phonological space that reflect the relative occurrence of certain features. Whereas most nouns and verbs are typicality neutral, containing form features that are equally common in the two categories, there are also clearly typical nouns and verbs (nouns that are more typical share fewer features with verbs and vice versa). Farmer et al. found that English speakers were faster to read typical words than atypical words.

Tanenhaus and Hare (2007) argued that the findings of Farmer et al. (2006) might help explain eye movement patterns during reading: The length of first fixations could be contingent upon form-feature predictions. This would be consistent with an early visual M100 effect for words containing unexpected form features. From this perspective, the visual M100 component would be sensitive to the probabilistic distribution of form features across the entire mental lexicon, rather than being specifically tuned to detecting a small set of closed-class morphemes.

To test whether, in the context of syntactic prediction, closed-class morphemes have a special status as category indicators, or whether form typicality can also serve as a category cue for the visual cortex, we examined the visual M100 effect for three types of nouns presented in expected or unexpected contexts in word-by-word reading: (a) bimorphemic nouns (i.e., words with a closed-class category-marking morpheme, such
as “farmer,” “princess,” and “artist”), (b) monomorphemic, typical nouns containing form properties that are indicative of the noun category (e.g., “movie,” “soda”), and (c) neutral nouns (i.e., those with no clear form bias toward either nouns or verbs). The bimorphemic nouns and the typical nouns we used were about equally indicative of the noun category. To manipulate syntactic context, we created sentences in which the critical noun was preceded either by an adjective (e.g., “the beautiful ___”), such that a noun would be highly expected, or by an adverb (e.g., “the beautifully ___”), which would render nouns unexpected and instead induce a strong expectation for a participle (e.g., “dressed”).

If word-category violations are detected during early visual processing exclusively on the basis of closed-class morphemes in the input, then only the bimorphemic nouns should show an M100 effect of expectedness. Alternatively, if form typicality is sufficient, then an M100 effect should be present for typical nouns as well. Neutral nouns should not elicit an M100 expectedness effect according to either hypothesis.

In addition to comparing the average M100 response to each noun type in each sentence context, we analyzed dipole waveforms for single-trial data. This allowed us to conduct a multiple regression analysis addressing whether the presence of a closed-class morpheme leads to an M100 effect independently of a word’s form typicality.
6.2 Methods

6.2.3 Participants

Fifteen healthy, right-handed subjects participated (6 female and 9 male; average age = 23 years). All had normal or corrected-to-normal vision and gave informed consent.

6.2.4 Materials

Forty bimorphemic, 40 typical monomorphemic, and 40 neutral monomorphemic nouns were presented to participants in both expected and unexpected contexts (e.g., “the beautiful princess was painted” vs. “the beautifully princess was painted”). Nouns were drawn from Farmer and his colleagues’ (2006) analysis of the CELEX corpus (Centre for Lexical Information; Baayen, Piepenbrock, and Gulikers, 1995). Farmer et al. calculated the phonological distance between each possible two-word comparison based on the number of overlapping and non-overlapping phonetic features. They obtained a typicality score for each word by subtracting its distance to all verbs from its distance to all nouns. Typicality scores for the nouns and verbs in CELEX ranged from –0.632 to 0.498, with more highly negative scores denoting a more noun like form, scores around 0 denoting neutrality, and more highly positive numbers denoting forms more typical of verbs. The typical nouns had a mean score of –.42 (SD = .08), whereas the neutral nouns had forms that were approximately equally similar to nouns
and verbs ($M = .00$, $SD = .02$). The bimorphemic nouns were also typical of the noun category ($M = –.34$, $SD = .15$), but less so than the typical nouns. Target words were matched for frequency. Selecting words with suitable typicality values for our three noun types unfortunately resulted in length differences between conditions (neutral nouns were shortest, and bimorphemic nouns were longest). However, this did not appear to affect our results (see the multiple regression analysis in the Results section).

To avoid habituation, we used 240 matched filler sentences in which adjectives and adverbs were followed by participles (e.g., “the beautiful dressed . . .” or “the beautifully dressed . . .”).
6.2.5 Procedure

While seated in a dimly lit, magnetically sealed chamber, participants read the stimuli on a screen approximately 17 in. from their head. Sentences were presented word by word, with each word appearing for 300 ms, followed by a 300-ms blank screen. Participants judged each sentence’s grammaticality after the final word was presented. The entire MEG recording session lasted approximately 40 min. Data were collected using a whole-head 275-channel gradiometer (CTF, Vancouver, British Columbia, Canada) system sampling at 600 Hz in a band between 0.1 and 200 Hz. Further details regarding the procedure and materials can be found in the Supplemental Material available online.

6.2.6 Analysis

Data were high- and low-pass-filtered (at 1 and 40 Hz) and automatically cleaned of artifacts (approximately 10% of trials were rejected). To estimate the generating source of the M100, we used a multiple-source model (BESA Software; Brain Electrical Source Analysis 5.1, Gräfelfing, Germany) taking data from all sensors. Dipole locations did not differ across conditions, nor did the number of additional dipoles used in the model.
To test for M100 effects in the averaged data, we performed a 2 (context: expected vs. unexpected) × 3 (noun type: bimorphemic vs. typical vs. neutral) within-subjects analysis of variance (ANOVA) on the mean amplitude of a 15-ms interval centered around the average M100 peak for each condition and subject (as in Experiments 1 and 2). We used post hoc t tests to examine effects within each noun type.

To test for independent contributions of closed-class morphology and typicality to the M100 effect, we used an individual-trial mixed-effects regression analysis. We estimated peak M100 amplitude for each trial, using the previously generated source model, and then regressed amplitude against predictors for the effects of morphology, typicality, and other psycholinguistically relevant variables (listed in Table 1). To characterize how a word’s form typicality mismatches with the predicted form typicality, we estimated how far (in normalized units of typicality) the typicality of each encountered word lay from the mean typicality score of the expected word category. The coefficient for this regression term (predicted typicality mismatch) would be reliably greater than 0 if the difference between expected and encountered typicality affected the M100. To test whether closed-class morphology had a reliable independent effect, we included a Morpheme Presence (closed-class morpheme present vs. absent) × Context (expected vs. unexpected) interaction term.
6.3  Results

6.3.1  Averaged data: expectedness and M100 amplitude

Figure 1 shows the average M100 dipole activity in each condition. A 2 (context: noun expected vs. unexpected) \( \times \) 3 (noun type: bimorphemic vs. typical vs. neutral) within-subjects ANOVA on M100 amplitude revealed a main effect of context, \( F(1, 14) = 4.708, p = .048, \eta^2 = .252 \), and an interaction between context and noun type, \( F(2, 28) = 3.614, p = .017, \eta^2 = .467 \), indicating that this effect was not present in every condition. There was no main effect of noun type, \( F(1, 14) = 1.113, p = .299, \eta^2 = .169 \).

Figure 12 - M100 dipole waveforms per condition (Experiment 3)
Grand-averaged waveforms for the M100 dipole sources for expected and unexpected (a) bimorphemic, (b) typical, and (c) neutral nouns (n = 15). The dotted vertical lines indicate 15-ms intervals centered around the average M100 peak. Mean dipole locations and orientations (blue and red), as well as the dipoles from the individual participants (gray), are plotted for each noun type. Asterisks indicate a significant difference in M100 amplitude between unexpected and expected nouns (*p < .05).
Pair-wise comparisons confirmed that the difference in M100 amplitude between expected and unexpected nouns was reliable for the bimorphemic nouns, $t(14) = 4.18$, $p < .001$, $\eta^2 = .56$, but also for typical nouns, $t(14) = 2.15$, $p = .049$, $\eta^2 = .25$. Neutral nouns showed no effect, $t(14) = 0.32$, $p = .75$, $\eta^2 = .01$.

Because the latency of the M100 peak varied across subjects, we repeated the analysis using each individual’s by-condition peak amplitude as our dependent measure. This analysis produced essentially identical results, with reliable differences between expected and unexpected bimorphemic nouns, $t(14) = 3.634$, $p = .003$, $\eta^2 = .49$, and typical nouns, $t(14) = 3.171$, $p = .007$, $\eta^2 = .42$, but not neutral nouns, $t(14) = 0.733$, $p = .47$, $\eta^2 = .04$.

### 6.3.2 Single-trial analysis

The results of the regression are presented in Table 8. Despite the model’s high deviance score, indicating a low overall fit because of the noisy individual-trial data, the results are clearly interpretable. Controlling for all other variables, predicted typicality mismatch had a reliable effect on M100 amplitude, $\beta = 3.77$, $SE = 1.52$, $t = 2.49$, $p = .016$ ($p$ value simulated using Markov-chain Monte Carlo methods): The further a word’s form typicality was from the expected word category, the greater the M100 amplitude, a finding consistent with the results in the by-condition analysis.
However, the regression failed to provide any evidence that closed-class morphemes play a special role in generating an M100 effect. The increased M100 amplitude for unexpected nouns containing a closed-class morpheme was no greater than would be expected given their predicted typicality mismatch alone, as indicated by the small and non-significant interaction between the variables coding for context and morpheme presence (see Table 8).

One other reliable effect emerged from the regression: Nouns encountered in an unexpected context produced a reliably larger M100 than nouns encountered in an expected context, $\hat{\beta} = 2.84, SE = 1.36, t = 2.09, p = .04$ ($p$ value simulated using Markov-chain Monte Carlo methods). There was no effect of orthographic length, which suggests that the small length differences between conditions did not affect any of our results.
Table 8 - Linear regression results for single-trial M100 amplitude (Experiment 3)

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$\beta$</th>
<th>$SE$</th>
<th>$t$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>27.23</td>
<td>5.86</td>
<td>4.65</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>Context</td>
<td>2.84</td>
<td>1.36</td>
<td>2.09</td>
<td>.040</td>
</tr>
<tr>
<td>Morpheme presence</td>
<td>-1.96</td>
<td>1.57</td>
<td>1.25</td>
<td>.21</td>
</tr>
<tr>
<td>Predicted typicality mismatch ($M = 1.61$, $SD = 1.07$)</td>
<td>3.77</td>
<td>1.52</td>
<td>2.49</td>
<td>.016</td>
</tr>
<tr>
<td>Orthographic length ($M = 5.66$, $SD = 1.46$)</td>
<td>0.29</td>
<td>0.63</td>
<td>0.46</td>
<td>.64</td>
</tr>
<tr>
<td>Log frequency ($M = 5.85$, $SD = 1$)</td>
<td>0.25</td>
<td>0.47</td>
<td>0.53</td>
<td>.60</td>
</tr>
<tr>
<td>Number of syllables</td>
<td>0.62</td>
<td>2.37</td>
<td>0.26</td>
<td>.79</td>
</tr>
<tr>
<td>Orthographic neighborhood density ($M = 3.63$, $SD = 4.72$)</td>
<td>0.04</td>
<td>0.15</td>
<td>0.29</td>
<td>.77</td>
</tr>
<tr>
<td>Phonological length ($M = 5.48$, $SD = 1.36$)</td>
<td>0.18</td>
<td>0.75</td>
<td>0.24</td>
<td>.81</td>
</tr>
<tr>
<td>Morpheme Presence $\times$ Context interaction</td>
<td>0.46</td>
<td>1.98</td>
<td>0.23</td>
<td>.82</td>
</tr>
</tbody>
</table>

Note: The predictors listed were entered into the regression against peak M100 amplitude ($deviance = 29,354$; number of observations = 3,136). The $p$ values were simulated using Markov-chain Monte Carlo methods.

6.4 Discussion

When analyzed using both a factorial design and a multiple regression on individual trials, the MEG visual M100 response was sensitive to form typicality, and not just to a small set of closed-class morphemes. This strongly suggests that the brain uses prior syntactic context to predict not only a word’s syntactic category (e.g., Hale, 2006; Lau
et al., 2006; Levy, 2008), but also form features that are probabilistically associated with the predicted category.

A central aspect of any explanation of these occipital word-category effects is whether the effects arise in an entirely top-down fashion, or, alternatively, whether the regions generating the visual M100 house some type of category representations. Our results cannot strictly settle this issue, as it is impossible to discern whether the M100 effect we observed results from low-level form-feature matching or from a true word-category mismatch. Importantly however, these results open up the possibility that the M100 effect observed in Experiments 1 and 2 for unexpected participles was due to a ‘shallow’ form-feature mismatch, with [-ed] being an unlikely noun ending, instead of being due to a mismatch at the level of word category analysis. This would predict that the absence of an effect for the bare stems and unambiguous nouns was due to form typicality of the noun items that we used, which, given the distribution of English nouns on the typicality scale, was likely to average around 0 (i.e., neutral).

In fact, when subdividing the unambiguous nouns from Experiment 2 into one group of typical nouns (18 x 2 items) and a group of nouns whose typicality scores averaged around zero (17 x 2 items), we see that the M100 effect emerges for the typical subgroup, but not for the neutral group.
Figure 13 - M100 dipole waveforms for unambiguous nouns (Experiment 2)

Grand-averaged waveforms for the M100 dipole sources for expected and unexpected unambiguous nouns (from Experiment 2), binned into typical and neutral groups of nouns. The dotted vertical lines indicate 15-ms intervals centered around the average M100 peak. Asterisks indicate a significant difference in M100 amplitude between unexpected and expected nouns (*p < .05).

This suggests that form neutrality, and not the absence closed-class morphology, was the cause for the absence of an effect for the bare stems and unambiguous nouns in Experiments 1 and 2.

Before moving on to asking whether lexical-semantic violations may similarly trigger enhanced visual responses, two final points are in place.

First, it is important to note that effects of form typicality on reading times have not been robust. For example, Staub, Grant, Clifton, and Rayner (2009) failed to replicate Farmer et al.’s (2006) findings when using a mixed design rather than a
blocked design (noun and verb items in separate experiments; as utilized in the original experiments by Farmer et al. 2006). Based on this apparently dramatic influence of experimental design on the presence vs. absence of an effect of typicality, Staub et al. concluded that the effects found by Farmer et al. might “reflect task-dependent strategic factors as opposed to the processes involved in normal word recognition” (p. 813). However, this conclusion is not necessarily warranted. First, it is unclear what would motivate participants to engage in task-specific strategies in this particular case: Farmer et al. (2006) used as many typical as atypical nouns and verbs, rendering form typicality a terrible word category cue within the experimental environment. Further, Farmer and colleagues (subm.) recently carried out an exact replication of Staub et al.’s (2009) mixed design and found that reading time effects of phonological typicality reduced over time (present in the first half but not the second half), which is contrary to what one would expect if participants were developing experiment-specific strategies. Thus, it seems that, if anything, exposing participants to the full range of typicality scores has an attenuating effect: Whereas in ‘real life’ typicality scores range from neutral to typical, with a few exceptions (recall that most nouns and verbs are typicality-neutral), suddenly participants are confronted with a situation where atypical form features are just as likely as typical form features. This may be an example of the flexibility of the neural circuitry underlying sentence processing, especially in cases where properties are probabilistic in nature. This is not the only situation where experimental effects have been found to be attenuated during
the course of an experiment or as a result of the experimental manipulation. Anyone who has ever participated in a grammaticality or sensicality judgment task has experienced the sensation of not being able to judge even the simplest violations anymore after having been exposed to a number of ungrammatical or unlicensed constructions. In fact, it has been shown that the ratio of grammatical vs. ungrammatical constructions affects ERPs, in particular later-peaking responses. For example, in an experiment testing word category violations of the type given under (2) above, Hahne and Friederici (1999) found an ELAN effect followed by a P600 effect when participants were presented with a violation ratio of 50%, but no P600 effect when 80% of the sentences contained word category violations (the ELAN persisted).

That we still found an effect of unexpected form features throughout the entire experiment (in contrast to Staub et al., 2006 and Farmer et al., submitted) could thus be due to the fact that we did not use the full range of typicality scores and thus preserved the cue-validity of form typicality with respect to word category: While word category expectations were violated, this was never the case for the relationship between form features and word category (no atypical nouns were used).

The second point that I want to highlight is that, based on the present results, we can conclude that closed-class morphology is not necessary for the early detection of violations of word category expectations (and consequently that word category access need not take place during early visual analysis), but we cannot say anything definitive with regard to the exact nature of the form representations to which the
visual M100 generator responds. The form typicality scores used by Farmer et al. (2006) and in this experiment are based on phonological features (e.g., +/- voiced), and while there exists a level of orthographic transparency in English, it is not immediately obvious how phonological typicality is to be translated into the visual modality. Further, it is unclear whether form predictions are very fine-grained, e.g., down to the level of individual letter forms, or whether they might be operating on e.g., general shape of words, etc. Future research should be aimed at elucidating the nature of the form representations over which (mis)match effects are found during the M100 response, for example by way of disrupting potentially relevant levels of form representation (e.g., case alternation).
Chapter 7 - Experiment 4

7.1 Introduction

Experiment 3 demonstrated that the M100 component is sensitive to form properties of words as they are associated with syntactic category. In Experiment 4, I test whether the M100 component might be similarly sensitive to unexpected form features for violations of lexical-semantic predictions. As pointed out in the Introduction however, lexical-semantic violations studied in EEG typically affect amplitude of the N400 component, a long-lasting ERP response peaking around 400ms after stimulus onset (e.g., Kutas, Van Petten, and Kluender, 2006).

There are a number of possible reasons why no early sensory effects have been previously found for lexical-semantic violations.

Under one set of accounts, syntactic prediction and lexical-semantic prediction constitute qualitatively different processes. For example, one might hypothesize that syntactic prediction involves the top-down activation of form representations, whereas lexical-semantic prediction does not. However, previous studies suggest that lexical-semantic prediction does in fact include the preactivation of word form properties. For example, Laszlo and Federmeier (2009) show that overall orthographic similarity to predicted words affects N400 amplitude. Similar experiments in the auditory domain have demonstrated that words that violate phonological, but not semantic, predictions
generate an ERP effect that can be dissociated from the N400 response (the phonological mismatch negativity; see e.g., Connolly and Phillips, 1994). Based on their findings that orthographic similarity to predicted words affects the N400 response to incoming letter strings independent of whether these constitute real words, Laszlo and Federmeier (2009) have argued that lexical-semantic prediction poses a filter on the orthographic processing of incoming words prior to lexical access. Such an account bears similarity to the Sensory Hypothesis I proposed above syntactic prediction. One obvious difference between these studies and the studies discussed in Chapters 4-6, however, is that the N400 and the phonological-mismatch negativity both likely reflect later stages of processing than the MEG M100 response. This leads to the hypothesis that differences between lexical-semantic and syntactic prediction might not so much lie in the presence vs. absence of top-down processing, but rather in the nature and localization of the representations of relevant form properties. In other words, maybe the generator of the visual M100 response responds to syntactically relevant form-representations only. If this is the case, no M100 effect of lexical-semantic violations should be found under any circumstances.

A finding where the M100 generator selectively supports the representation of syntax-specific form-features would be quite surprising, especially in light of the fact that the visual M100 response to visual objects and words in isolation shows no selectivity for stimulus-type (not distinguishing between letters and symbols, Tarkiainen et al., 1999). But it would be a very interesting result, suggesting that
syntactic and lexical-semantic processing are distinct from the earliest stages of visual analysis, i.e., the encapsulation of these two processes would extend beyond areas in left-temporal and prefrontal brain regions.

Under another explanation for the absence of early sensory effects for lexical-semantic violations, lexical-semantic and syntactic prediction are not inherently distinct processes. Rather, the fact that no early sensory effects have been previously reported for lexical-semantic violations could be a by-product of differences regarding the relationship between syntactic category and word form on the one hand, and lexical-semantic field and word form on the other. While for syntactic category a clearly definable form correlate has been established (form typicality, see Chapter 6) no such straightforward form correlate has been identified for lexical-semantic fields. There exist some examples of words that are related in both form and meaning (such as boil/broil), but form similarity and meaning similarity typically do not go hand in hand. For example, while rose and magnolia are semantically related but unrelated in form, rose and hose show form, but not meaning similarity. As a result, lexical-semantic predictions might be translated into form-based estimates only when there is either a single lexical element predicted (a guaranteed one-to-one mapping between meaning and form) or when form features are sufficiently shared between multiple candidates. In the majority of N400 studies, however, context does not predict 100% for a particular word for all items in the experiment. Similarly, while factors like bigram frequency and orthographic and phonological neighborhood density are
typically taken into account in studies that investigate lexical-semantic processing, it is not usually asked if the violating word is sufficiently distinct from the predicted word in terms of its form features.

A further possible explanation for the absence of an early sensory effect for lexical-semantic violations is methodological in nature. Most studies investigating lexical-semantic violations use EEG, for which analyses are usually performed over averages across relatively large time-intervals. In contrast to the N400 however, sensory components like the M100 do not constitute long-lasting sustained activity, and thus a peak-focused analysis of amplitude variation as a function of predictability might be more appropriate. In addition, early sensory components are not typically the focus of analysis in studies investigating lexical-semantic processing, so early effects might be either overlooked or not interpreted.

One recent EEG study on lexical-semantic violations did find effects peaking before 100ms post-stimulus onset over left-posterior electrodes for predicted vs. unpredicted words given prior discourse (Dambacher et al., 2009). Dambacher and colleagues argue that SOA (Stimulus-Onset Asynchrony) might have been a crucial factor in explaining the difference between their findings and earlier studies: In their study, words were presented for 250ms separated by 50ms blank screens. However, we found early sensory effects for syntactic violations with a ‘slow’ word-by-word reading paradigm (300ms on/300ms off), which casts doubt upon long SOA’s being
the only explanation for the absence of early lexical-semantic effects in the majority of studies.

The present study was designed to ask whether the M100 response can be sensitive to violations of lexical-semantic predictions. To ensure that a mapping between form and meaning could be drawn, I employed a simple picture-noun matching task instead of studying lexical-semantic violations using high cloze-probability sentential contexts. Pictures were used to set up specific expectations for just one lexical item in the subsequent noun phrase and participants knew beforehand which word was the appropriate match. In the case of a violation of a prediction, the target word’s form was manipulated such that it was maximally distinct from the predicted word. Although this design raises the question whether the task might be too far divorced from natural language processing, it does potentially set up optimal conditions for obtaining early sensory effects.

I further included a match/mismatch comparison where no specific prediction for a particular word form could be generated based on the preceding picture. Instead, the picture presented in these conditions represented an entire semantic field (animals and food/drink items respectively). For this manipulation, no early visual effects were expected.

In sum, in Experiment 4 I varied the nature of the picture prime (context: +predictive vs. −predictive), and whether or not the noun phrase matched the preceding picture (congruence: match vs. mismatch).
To assess whether early visual effects might be obtained for violations of predictions for specific lexical entries, I performed a targeted analysis on the M100 component, focusing on peak activity instead of averaging across long time-intervals.

Later time-windows

Recall that I suggested based on the findings from Experiment 2 that no lexical access takes place during the M100 response (no sensory effect was found for unexpected word category unambiguous nouns like tree). Therefore, instead of an M100 effect, in the present case we might find correlates of cost associated with accessing lexical representations for matching vs. non-matching nouns in later time-windows. In fact, many observations of amplitude reductions of the N400 component for highly predictive contexts can be explained in terms of a lexical access account (e.g., DeLong et al., 2005; Federmeier and Kutas, 1999; Federmeier, 2007; Lazlo and Federmeier, 2009; Lau, Almeida, Hines and Poeppel, 2009; Lau, Phillips, and Poeppel, 2008). In MEG, the M350 response, a component peaking around 350ms post-stimulus onset localized to left-temporal cortex, likely reflecting a subcomponent of the underlying generators of the N400 response (Pylkkänen and Marantz, 2003), has been shown to be affected by lexical factors for words in isolation (Pylkkänen, Stringfellow, and Marantz, 2002) as well as semantically anomalous sentences (Pylkkänen and McElree, 2007). Lau et al. (2009) recently directly compared lexical priming effects for words presented in isolation and semantic (in)congruence effects for words presented in...
sentence contexts in MEG. They found that priming effects and contextual compatibility effects were qualitatively similar, which supports a view wherein at least part of the N400 component reflects lexical access.

Thus, in the present study we might expect reduced brain activity in the N400-time-window for nouns whose lexical entries have been primed by the preceding picture (i.e., by way of a picture denoting a specific object) compared to nouns that have not been contextually primed. For pictures that represent semantic fields (food or animals) it is not immediately obvious what to expect. When participants see e.g., a picture of a grocery bag, they might start generating potential matches, activating a few or even many lexical entries corresponding to food items. One would then expect reduced brain activity in response to words that belong to the food/drink category as opposed to the animal category, but only if the presented noun happens to be one of the lexical entries that were pre-activated either in a targeted way or via spreading activation (e.g., Federmeier and Kutas, 1999). However, it may be unlikely that pre-activated and encountered lexical representations coincide across all trials, and we might get a null effect, i.e., no differences between matches and mismatches, instead. A null effect would also be expected if no pre-activation of lexical entries occurs at all based on pictures representing semantic fields.
Task-dependent factors

Since picture-noun matching is a very explicit and controlled task compared to natural language processing, it is not unlikely that participants will employ task-specific strategies, or develop these in the course of the experiment. One possibility is that participants engage in different strategies depending on the nature of the picture prime.

If picture types are randomly intermixed, this might then mean that participants will switch task-strategies between predictive and non-predictive trials. Task-switching cost effects on both behavioral and neural responses have been well-established, especially when tasks are mixed within blocks (see e.g., Jamadar et al., 2010; Monsell, 2003 for a review). It has also been suggested that participants may not successfully perform task-switches across all trials (De Jong, 2000). As a result, participants might treat predictive contexts as non-predictive contexts on a number of trials. If it is true that prediction is crucial in enabling early sensory effects, it is important that conditions for predictive processing are optimal. Thus, to avoid any risk of task-switching from interfering with optimal predictive processing, we presented participants with a block of trials with only specific picture-noun pairs in addition to a block where specific pictures and pictures representing semantic fields were intermixed. If task-switching were to wash away any early prediction effect, we might expect these effects to only emerge during the non-intermixed block.

Alternatively, early visual responses to lexical-semantic violations might emerge over time resulting from a learning effect, in which case we would only expect
an M100 effect in the second half of the experiment. In other words, while effects at the level of form features may disappear if their cue-validity is reduced or removed (see previous chapter), early visual effects of lexical-semantic predictions may emerge under conditions where the relationship between lexical-semantic and form properties is enhanced.

In sum, the primary goal of this experiment was to establish whether violations of lexical-semantic predictions can affect early sensory processing, just like violations of syntactic predictions. Such a finding would strongly challenge a strict bottom-up view of language processing and would support models where contextual predictions facilitate top-down effects on low-level sensory regions.

7.2 Methods

7.2.1 Participants

After the exclusion of 3 participants based on low signal-to-noise ratio in the MEG data, and 4 other participants because they did not meet the criteria of showing typical M100 and M170 field patterns (see below), 15 participants were included in the analysis below.
7.2.2 Materials

80 different nouns were presented in four conditions in a 2 (Context: +Predictive vs. – Predictive) by 2 (Congruence: Match vs. Mismatch) design. Noun phrases either followed a picture denoting a specific object (+Predictive; 40 animals and 40 food/drink items) or they followed a picture representing a semantic field (- Predictive). Nouns either matched the image (match) or they did not (mismatch).

As mentioned above, the function of the pictures was to constrain cloze-probability. For example, a picture of an apple was used to induce a strong prediction for the word *apple* embedded in a subsequent noun phrase. A noun such as *carrot* then violated this strong prediction. Since it is not clear which word form properties exactly are processed during the visual M100 stage (general word shape, letter sequences, letter presence, or ‘form features’), visual properties of the words were manipulated such that predicted and encountered words in the mismatch condition were maximally distinct in terms of Levenshtein edit-distance (the number of letters that need to be replaced, removed, or added to convert one word into another; e.g., Navarro, 2001) and word length. To achieve the latter, nouns were binned into two groups of ‘short’ (3-5 letters) and ‘long’ (5-8) nouns. For the mismatch condition then, if a ‘short’ noun was expected, a long noun was presented instead, and vice versa.

The –Predictive conditions also consisted of picture-NP sequences, but instead of specific objects, images referred to semantic fields, either animals or food/drink
items: A picture of a grocery bag stood for any word describing an edible or drinkable object, and a picture denoting Noah’s Ark functioned as a placeholder for any noun describing an animal. For example, a grocery bag followed by the milk would constitute a match while a noun phrase like the dog was a mismatch, and the other way around if these noun phrases followed a picture of Noah’s Ark (80 per condition; 40 animals; 40 food/drink items). As pointed out in the Introduction, participants could in theory be anticipating or pre-generating a number of words at presentation of e.g., the grocery bag. However, it is unlikely that participants were activating the word form for a single word or a for few words that share form properties. In other words, -Predictive here refers to the assumed absence of a word form prediction—based on a direct mapping between meaning and form—that is narrow enough to generate an early visual mismatch effect if this prediction is violated.

Both context and target words were maximally balanced: Input preceding the target noun was kept constant across conditions for 1500 ms (from the offset of the picture to the onset of the target noun; see Figure 14) and each noun (40 nouns denoting animals and 40 nouns denoting food/drink items) appeared once in each condition, e.g., the word apple appeared once as a matching noun in the +Predictive/Match case, once as a mismatching noun in the +Predictive/Mismatch condition, once as a matching noun in the –Prediction/Match condition, and once as a mismatching noun in the –Prediction/Mismatch condition. The design is presented in Table 9.
Table 9 - Examples of experimental stimuli (Experiment 4)

<table>
<thead>
<tr>
<th>+ / - PRED</th>
<th>PRIME</th>
<th>TARGET</th>
<th>+ / - MATCH</th>
<th># of items</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[any food/drink item]</td>
<td>the apple</td>
<td>+ MATCH</td>
<td>BLOCK 1: 80</td>
</tr>
<tr>
<td></td>
<td>[any animal]</td>
<td>the apple</td>
<td>+ MATCH</td>
<td>BLOCK 2: 80</td>
</tr>
<tr>
<td>+ prediction for specific word (form)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[any food/drink item]</td>
<td>the apple</td>
<td>+ MATCH</td>
<td>BLOCK 1: 80</td>
</tr>
<tr>
<td></td>
<td>[any animal]</td>
<td>the apple</td>
<td>+ MATCH</td>
<td>BLOCK 2: 80</td>
</tr>
<tr>
<td>- prediction for specific word (form)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[any food/drink item]</td>
<td>the apple</td>
<td>- MATCH</td>
<td>BLOCK 2: 80</td>
</tr>
<tr>
<td></td>
<td>[any animal]</td>
<td>the apple</td>
<td>- MATCH</td>
<td>BLOCK 2: 80</td>
</tr>
</tbody>
</table>

* BLOCK 1: only + PRED conditions (80 x match / 80 x mismatch)
* BLOCK 2: + PRED conditions intermixed with - PRED conditions

Examples of the experimental stimuli of Experiment 4 (80 per condition; 40 animals and 40 food/drink items). +Predictive conditions were presented twice: once in isolation (Block 1) and once intermixed with –Predictive conditions (Block 2)

Figure 14 - Stimulus presentation (Experiment 4)

Pictures were presented for a duration of 900 ms, followed by 300 ms blank screen, after which a fixation cross appeared and then the noun phrase (300 ms on/off). A question mark prompted participants to indicate whether or not the noun phrase described the preceding picture.
7.2.3 Procedure

During the experiment, participants lay in a dimly lit, magnetically shielded room. Stimuli were presented using PsyScope X, and projected onto a screen positioned approximately 50cm from the participant’s eyes. Pictures were presented enclosed by a white box on a 75% grey background. Words were presented in white in non-proportional Courier font (size 28) against a 75% grey background.

Each trial began with the presentation of a picture (see Materials) for 900ms, followed by a 300ms blank screen, after which a fixation cross appeared and then the noun phrase (word-by-word, 300 ms on / off). A question mark followed the noun phrase, prompting participants to press the MATCH button (left index finger) if the noun phrase accurately described the preceding picture, or the MISMATCH button (left middle finger) if the noun phrase did not match the preceding image. At button press, the question mark was replaced by a feedback screen (giving the correct answer “MATCH” or “MISMATCH”). Appearance of the next trial was self-paced, and subjects were told that they could take breaks during feedback screens if they needed to. The experiment was divided into 2 blocks, separated by a screen letting participants know how far along in the experiment they were. Block 1 consisted only of +Predictive trials (see Introduction; presented in random order; 160 trials total). Block 2 contained both +Predictive and –Predictive trials (intermixed in random order; 320 trials total). The +Predictive trials were an exact repetition of the trials that had been
presented in Block 1 (but not in the same order). The experiment was preceded by training and practice. Participants were familiarized with the task (15 practice trials), and it was explained to them that a picture of a grocery bag functioned as a stand-in for “any food(drink item” and that a picture of Noah’s Ark represented “any animal.” Under the hypothesis that an M100 effect will only be seen for cases where a prediction for a single lexical item is violated, it is important to ensure that participants indeed associate a single word with each picture. Therefore, prior to the experiment participants were shown all objects that were used as picture primes with the corresponding noun written below each image. They were asked to look at the image and read each word before proceeding to the next picture. Another 15 practice trials were given after the participant was seated in the magnetically shielded room, right before the experimental trials were initiated.

Neuromagnetic fields were recorded continuously with a whole-head, 157-channel axial gradiometer array at a sampling rate of 1000 Hz in a band between 0 and 200 Hz. A notch filter of 60 Hz was applied. The entire recording session lasted approximately 45 minutes.

7.2.4 Analysis

Pre-processing

Prior to averaging, the MEG data were cleaned of artifacts (in BESA 5.1) and trials on which participants provided an incorrect judgment. On average, this resulted in the
exclusion of 15% of the data per subject. Data were averaged by stimulus category over a 900ms epoch with a 300ms pre-stimulus interval, time-locked to the appearance of the target word. Prior to analysis, recordings were high and low-pass filtered at 1 and 40 Hz respectively (similar to Experiments 1-3).

Sensors-of-interest analysis: Visual M100 and Visual M170

As a first step, each participant’s grand-average was inspected for the presence of typical field patterns associated with the visual M100 and the visual M170 components (for a description, see Section 2.3). 4 participants did not show both patterns and were thus excluded from the analysis. This resulted in the inclusion of 15 participants in the analyses reported below.

For both the visual M100 and the visual M170 components, sensors of interest were selected based on grand-average data across all participants and conditions (n = 15). For each component, the time point of peak activity was determined via visual inspection (M100 peak time: 97 ms post-stimulus onset; M170 peak time: 151 ms post-stimulus onset). Based on the magnetic field pattern of each component’s grandaveraged peak, sensors were identified that showed amplitudes that diverged more than 40fT from zero for the M100 peak and more than 45fT for the M170 peak.

As can be seen in Figure 15A, the re-entering field of the M100 component (blue/left-hemisphere) showed more activity than the outgoing field (red; right-
hemisphere). A similar hemispheric difference can be seen for the M170 response in Figure 15B (but note the opposite directionality). This asymmetry replicates the results presented for the sensor data analysis in Experiment 1, and could reflect that the left-hemisphere sources might be more active than the right-hemisphere sources underlying the M100 and M170 components. However, as already mentioned in Chapter 4, caution with this conclusion is warranted (for the M100 response in particular), since the exact relationship between the bilateral underlying sources and the single field pattern that reaches the skull surface cannot be inferred.

Regardless, only very few right hemisphere sensors met the threshold levels employed for the sensor selection (40fT for the M100 component and 45fT for the M170 response), and all analyses reported below were performed over the left-hemisphere fields’ sensor groups only (although it is worth noting that there were no differences in the pattern of results as a function of hemisphere).

For each component, sensor waveforms for the sensors-of-interest were extracted for each condition and subject separately. Analyses were conducted over activity across a 10ms interval of interest, centered around the peak activity, defined for each condition separately (indicated by the vertical dotted lines in Figures 18 and 19).

These values were then entered into two separate within-subjects ANOVA’s. The first ANOVA, which included only stimuli presented in Block 2, was aimed at investigating whether an M100 effect, if at all present, is dependant on whether or not
a specific word could be predicted based on the picture (2 (Context: -Predictive vs. +Predictive [Block 2]) by 2 (Congruence: Match vs. Mismatch).

A second 2 (Block: Block 1 vs. Block 2) by 2 (Congruence: Match vs. Mismatch) within-subjects ANOVA aimed at estimating whether there were reliable differences between Block 1 and Block 2 with respect to any effect of Congruence, including +Predictive contexts only. ANOVA’s were conducted for the M100 and M170 components separately.

**Sensors-of-Interest analysis of later components**

As pointed out in the Introduction, I also aimed to investigate effects of prediction/integration on later components that fall in the N400 time-window. Figure 15C and 15D show the field pattern for two components peaking around 300 ms and 450 ms respectively, with outgoing and re-entering fields concentrated over left- and right-anterior sensors.

It has been previously noted that the long-lasting N400 response probably reflects a complex of several sub-components (Pylkkänen and Marantz, 2003; Van den Brink, Brown, and Hagoort, 2001; Halgren et al., 2002), and a recent MEG study by Lau et al. (2009) similarly report two components with distinct field patterns within the N400 time-range.
The second component in particular (15D) shows a field pattern that suggests it might reflect the M350 component, with two bilateral temporal sources. The first component (15C), exhibits a field pattern that may reflect a ventro-medial source that has previously been associated with (semantic) composition (Pylkkänen and McElree, 2007; Brennan and Pylkkänen, 2008; 2010; Pylkkänen, Oliveri, and Smart, 2009; Pylkkänen, Martin, McElree, and Smart, 2009; Bemis and Pylkkänen, submitted).

Since both components were long-lasting and corresponding magnetic field patterns underwent slight changes over time, peak-based extraction of sensors-of-interest was not entirely appropriate. Instead, we conducted the analysis over the two anterior quadrants of sensors. As can be seen in Figure 19, both left- and right-anterior sensors showed the bi-phasic pattern predicted by the field patterns given in Figure 15, with one sustained component lasting from ~ 200 to 400 ms and another from ~ 400 to 600 ms.

To investigate effects of congruence at each component, sensor data was averaged by condition across two 200 ms time-intervals (200-400ms and 400-600ms) and entered into two separate ANOVA’s with the same factors and levels as described above for the M100 and M170 components: A 2 (Context: -Predictive vs. +Predictive [Block 2]) by 2 (Congruence: Match vs. Mismatch) within-subjects ANOVA investigated whether there were any effects of congruence for stimuli presented in Block 2, and a 2 (Block: Block 1 vs. Block 2) by 2 (Congruence: Match vs. Mismatch) within-subjects ANOVA aimed to assess whether there were any differences between
the two blocks with respect to congruence effects for the +Predictive comparisons. Amplitudes over sensors pertaining to re-entering fields were entered as positive values in all analyses.

Figure 15 shows the field patterns and average peak time of each component that was analyzed. Sensors of interest are marked in black.

**Figure 15 - Grandaveraged field patterns for each component of interest (Experiment 4)**

Grandaveraged magnetic field distributions for all participants (n=15) and conditions combined, of the M100 component (15A), M170 component (15B), and two later, more sustained components (15C and 15D respectively). Sensors-of-interest over which analysis was performed are marked in black on the brains displaying the field distributions.(plotted in grey next to the component of interest).
7.3 Results

7.3.4 Behavioral Results

Error rates

Participants showed an overall average accuracy of 97% (SD = .8%). Regardless however of this low overall error rate, there were reliable differences between conditions with respect to accuracy.

A 2 (Context: -Predictive vs. +Predictive [Block 2]) by 2 (Congruence: Match vs. Mismatch) within-subjects ANOVA over % correct trials per condition across Block 2 showed a main effect of Context ($F(1,14) = 11.742, p = .004$): participants committed less errors overall for the +Predictive conditions (M = 1.29 %, SD = .4) than the –Predictive conditions (M = 5.96 %, SD = 1.65). Further, there was a main effect of Congruence ($F(1,14) = 7.226, p = .018$), but also a Context x Congruence interaction ($F(1,14) = 8.568, p = .011$): While participants performed significantly better on mismatch trials for the –Predictive conditions ($t(14) = 2.981, p = .010$), this was not the case for the +Predictive cases ($t(14) = .174, p = .864$).

Recall that all +Predictive trials presented in Block 2 had been presented to the participants already in Block 1. However, a 2 (Block: Block 1 vs. Block 2) by 2 (Congruence: Match vs. Mismatch) within-subjects ANOVA over +Predictive trials
did not show any effects of Block, but this might be due to the fact that accuracy was already at ceiling in Block 1.

**Response Times**

A 2 (Context: -Predictive vs. +Predictive [Block 2]) by 2 (Congruence: Match vs. Mismatch) within-subjects ANOVA showed a main effect of Context ($F(1,14) = 56.743$, $p < .001$) and a main effect of Congruence ($F(1,14) = 9.452$, $p = .008$), indicating that participants were faster overall for the +Predictive conditions than the -Predictive conditions, and for the Match than the Mismatch conditions. There was no interaction between Context and Congruence ($F(1,14) = .944$, $p = .348$). However, planned comparisons showed that there was a reliable effect of Congruence for the -Predictive contexts ($t(14) = 2.364$, $p = .033$) but not for the +Predictive trials ($t(14) = .859$, $p = .404$).

A 2 (Block: Block 1 vs. Block 2) by 2 (Congruence: Match vs. Mismatch) within-subjects ANOVA over Response Times revealed a main effect of Block ($F(1,14) = 13.426$, $p < .003$): participants were significantly faster for Block 2 than for Block 1. No other effects were found.
**Summary Behavioral Results**

In sum, participants performed better and faster for the +Predictive conditions than for the –Predictive conditions. Effects of congruence were only found for the –Predictive contrast.

Further, participants were faster for the +Predictive conditions in Block 2 than in Block 1. Note that this direction is opposite to the behavioral slow-down that would be predicted under an account where task-switching (within Block 2) between predictive (forward) and non-predictive (backward) matching is costly. However, it is possible that facilitation of trial repetition outweighed any effect of task-switching.

A repetition-related interpretation of the results would predict an immediate speed-up at the start of Block 2 (as a task-switching account would predict an immediate slow-down). However, Figure 17, which plots response times across trials for the +Predictive condition (collapsed over Match and Mismatch trials), clearly shows that the speed-up is non-linear: Participants get faster over time during Block 1 and into Block 2, until reaching ceiling half-way through Block 2. Most importantly, the fact that there is no sharp division line between the two blocks (the ceiling effect is not observed until about 1/3 into the second block), and that participants start improving from the beginning of the experiment, suggests that neither task-switching nor stimulus repetition had a categorical effect on response time.
131

Figure 16 - Trial-by-trial response times (Experiment 4)
Response times for the +Predictive conditions (Match and Mismatch collapsed) to individual trials for all participants (320 x 15 trials excluding inaccurate responses as well as 2 SD outliers: 2696 trials), in order of presentation. The transition point between Block 1 and Block 2 is indicated by the vertical dotted line.

Although general improvement in response times across the experiment could suggest that participants were getting better at timing their response to the offset of the noun, this pattern of results raises the question if participants might have been improving in predictive processing. If this is the case, we might expect that sensory effects emerge over time as well, and we would predict that an amplitude effect on the M100 component is more likely to surface in Block 2 than Block 1.

7.3.2 MEG results: M100 sensors

Figure 18 displays the results of the sensors-of-interest analysis on the M100 component described above. Recall that amplitude over M100 sensors was extracted
from the re-entering (left-hemisphere) field of the M100 component, here plotted as negative values. Average peak time-points per condition are marked by the vertical dotted lines. Stars indicate wherever differences were reliable (p < .05).

Figure 17 - M100 sensor waveforms per condition (Experiment 4)
Grandaveraged waveforms for the M100 sensors-of-interest per comparison (n=15; blue: match; red: mismatch). Peak M100 time-points (negative values) for each condition are indicated with vertical dotted lines. Sensors-of-interest are plotted in Figure 17D. * indicate wherever differences were reliable at p < .05.
**Match vs. Mismatch**

To investigate whether the M100 component was sensitive to Congruence, peak amplitude was entered in a 2 (Context: -Predictive vs. +Predictive [Block 2]) by 2 (Congruence: Match vs. Mismatch) within-subjects ANOVA (only including trials from Block 2). This showed a marginal effect of Congruence ($F(1,14) = 4.097, p = .062$), no main effect of Context ($F(1,14) = .001, p = .972$), and no interaction between Context x Congruence ($F(1,14) = 1.845, p = .196$).

However, post-hoc t-tests showed that while there was no effect of Congruence in the -Predictive trials ($t(14) = .058, p = .95$; Figure 17C), M100 amplitude over Mismatch trials was significantly increased compared to Match trials in the +Predictive trials ($t(14) = 2.358, p = .033$; Figure 17B). This result corroborates the hypothesis that the visual M100 component is sensitive to whether a word matches a lexical-semantic prediction, but only if there is a straightforward form correlate of the lexical-semantic level (in the +Predictive manipulation, but not the -Predictive manipulation).

**Block 1 vs. Block 2**

For response time, we saw that subjects performed faster for the +Predictive trials in Block 2 than they did in Block 1. To investigate whether there was also a difference in the M100 effect for the first vs. second presentation of the same trials (Block 1 vs. Block 2), M100 peak amplitudes for +Predictive trials were entered in a 2 (Block:
Block 1 vs. Block) by 2 (Congruence: Match vs. Mismatch) within-subjects ANOVA. This analysis showed a main effect of Congruence ($F(1,14) = 9.136, p = .009$), but no main effect of Block ($F(1,14) = .006, p = .942$). There was also no interaction between Block x Congruence ($F(1,14) = .540, p = .475$), but planned comparisons did reveal that, while the M100 effect of Congruence was reliable for Block 2 ($t(14) = 2.358, p = .033$; Figure 17B), the M100 amplitude difference between match and mismatch trials was not (yet) reliable in Block 1 ($t(14) = .933, p = .366$; Figure 17A). However, as can be seen in the bottom panel of Figure 18, it seems that the M100 effect was already emerging during Block 1 (compare M100 amplitude differences in the –Predictive trials to differences in the +Predictive (Block 1) trials).

This result tentatively supports the possibility raised in the discussion of the behavioral results above that participants might have become ‘better predictors’ as the experiment progresses.

M100 amplitude was not affected by word length.

### 7.3.3 MEG results: M170 sensors

2 (Context: -Predictive vs. +Predictive [Block 2]) by 2 (Congruence: Match vs. Mismatch), and 2 (Block: Block 1 vs. Block) by 2 (Congruence: Match vs. Mismatch) within-subjects ANOVA’s did not reveal any effects at the M170 component, as shown in Figure 18. Recall that we also failed to show an effect of violations of form-
predictions on M170 amplitude in Experiment 1 (discussed in Chapter 4). This null result is thus compatible with the hypothesis that violations of form-predictions should affect similar brain responses for syntactic and lexical-semantic predictions. Note that the M170 sensors appear to pick up the M100 effect observed in Figure 18 (especially see Figure 18C). This is not surprising, given the large overlap between the sensors selected for the M100 response and the M170 response (see Figure 15, and compare Figures 17D and 18D). Recall, however, that the field patterns for both components are of opposite polarity and thus constitute clearly separable components.

Figure 18 - M170 sensor waveforms per condition (Experiment 4)
Grandaveraged waveforms for the M170 sensors-of-interest per comparison (n=15; blue: match; red: mismatch). Peak M170 time-points (negative values) for each condition are indicated with vertical dotted lines. Sensors-of-interest are plotted in Figure 18D. None of the three comparisons showed reliable differences.
7.3.4 MEG results: later components

Figure 19 displays the averaged sensor data per condition for the left- and right-anterior quadrants. These figures clearly show two long-lasting components, one from ~ 200 – 400 ms and one lasting from ~ 400 – 600 ms.

**Figure 19 - LA and RA sensor waveforms per condition (Experiment 4)**

Grandaveraged waveforms for Left-Anterior (top-panel) and Right-Anterior (bottom panel) sensors-of-interest per comparison (n=15; blue: match; red: mismatch). Sensors-of-interest are plotted in Figure 19D. * indicate wherever differences were reliable at p < .05.
Match vs. Mismatch in Block 2

200-400 ms time-window

Performed over trials presented in Block 2, a 2 (Hemisphere: Left vs. Right) by 2 (Context: -Predictive vs. +Predictive [Block 2]) by 2 (Congruence: Match vs. Mismatch) within-subjects ANOVA over sensor data from left- and right anterior quadrants averaged across the 200-400 ms time-window revealed a three-way interaction (Hemisphere x Context x Congruence; $F(1,14) = 5.133, p = .039$). There was a main effect of Congruence ($F(1,14) = 14.108, p = .002$) and a main effect of Hemisphere ($F(1,14) = 8.362, p = .011$), indicating that there was more activity overall across the left- than the right-anterior quadrant. However, Hemisphere did not interact with Congruence ($F(1,14) = 2.832, p = .114$): Sensors in both hemisphere were equally sensitive to whether a noun did or did not match the preceding picture.

In addition, this analysis showed a main effect of Context ($F(1,14) = 16.490, p = .001$): more activity was present for the –Predictive trials. This main effect is likely a result of an interaction between Context and Congruence ($F(1,14) = 7.339, p = .016$): As can be seen from Figure 19, for both the left- and right-anterior quadrants amplitude of the +Predictive/Match condition (Figure 19C/D) was reduced compared to the other three conditions (Figure 19C/D vs. 19E/F). Planned comparisons for the left-anterior quadrant confirmed that there was no effect of Congruence for the -Predictive comparison ($t(14) = .499, p = .629$), while there was a reliable difference between Match and Mismatch trials for the +Predictive comparison ($t(14) = 4.439, p <
The right-anterior quadrant showed the same pattern (\( -\) Predictive: Match vs. Mismatch: \( t(14) = .333, p = .743; +\) Predictive: Match vs. Mismatch: \( t(14) = 6.450, p < .001 \)).

In summary, for the 200-400 ms time-window those trials where a unique noun was expected and presented, there was a significant reduction compared to all other conditions.

400-600ms time-window

A 2 (Hemisphere: Left vs. Right) by 2 (Condition: -Predictive vs. +Predictive [Block 2]) by 2 (Congruence: Match vs. Mismatch) within-subjects ANOVA over sensor data from left- and right anterior quadrants averaged across the 400-600 ms time-window (again performed only for trials presented in the second block) also showed a three-way interaction (Hemisphere x Context x Congruence: \( F(1,14) = 5.80, p = .030, \eta^2 = .293 \)). In contrast to the 200-400 ms window, there was no main effect of Hemisphere \( (F(1,14) = 1.970, p = .182) \). There was a main effect of Congruence \( (F(1,14) = 5.612, p = .032) \), with less activity for the Match conditions, as well as an interaction between Hemisphere and Congruence \( (F(1,14) = 6.016, p = .028) \). As can be seen in Figure 19, differences between match and mismatch trials observed over the left-anterior quadrant were not present for the right-anterior quadrant. There was no main effect of Context \( (F(1,14) = .490, p = .496) \) and the interaction between Context and Congruence was marginal \( (F(1,14) = 5.798, p = .087) \).
Planned comparisons for the left-anterior quadrant showed a similar pattern of results as for the 200-400 ms time-window: There was no effect of Congruence for the -Predictive manipulation ($t(14) = .402, p = .694$; Figure 19E), but a significant decrease for Match vs. Mismatch trials in the +Predictive comparison ($t(14) = 3.924, p = .002$; Figure 19C). For the right-anterior quadrant neither comparison was reliable (-Predictive: $t(14) = .093, p = .928$, Figure 19F; +Predictive: $t(14) = .771, p = .453$, Figure 19D).

In sum, as for the 200-400ms component, the 400-600ms component was modulated by Congruence for the +Predictive comparison only. In contrast to the 200-400 ms window, this effect seemed to be left-lateralized in this late time-window.

Block 1 vs. Block 2 for +Predictive trials

To assess whether effects in these later time-windows differed between the first and second presentation of the +Predictive trials (Block 1 vs. Block 2), we ran a 2 (Hemisphere: Left vs. Right) by 2 (Block: Block 1 vs. Block 2) by 2 (Congruence: Match vs. Mismatch) within-subjects ANOVA over both the 200-400ms time window and 400-600 ms time window. These ANOVAs showed no main effects of Block for either window (200-400ms: $F(1,14) = .012, p = .914$; 400-600ms: $F(1,14) = 1.568, p = .230$), and no interactions between Block and Congruence (200-400ms: $F(1,14) = 2.712, p = .415$; 400-600ms: $F(1,14) = .030, p = .865$).
In sum, while we saw that the M100 effect appeared to emerge only in Block 2, the later effects were already present in Block 1, and did not appear to attenuate or increase in Block 2.

7.4 **Discussion**

Using a picture-noun matching task, this study investigated whether early sensory effects can be found for lexical-semantic violations. An effect on amplitude of the M100 component was indeed observed, but only for cases where one particular word (one specific combination of form features) could be predicted. These findings are compatible with the hypothesis that the M100 response is sensitive to prediction-based form-estimates, as was suggested by the results of Experiment 3.

In contrast to Experiments 1-3, a reliable M100 effect only emerged about half-way through the experiment in this study, as evidenced by the fact that amplitude differences were only significant across the second block of the experiment. The second block differed from the first block in two respects: intermixing with non-predictive contexts and repetition of trials presented in the first block. However, neither of these factors can readily explain the pattern of results. If intermixing trial types would have interfered with predictive processing, we should have found effects in the first block and not in the second. If, on the other hand, repetition was the crucial factor, we should have seen a categorical distinction between the two blocks, with no
hint of an effect at all for the first block. But we did already see higher M100 amplitudes for mismatching nouns compared to matching nouns in the first block, although this difference was not reliable. Moreover, effects in later time windows did not differ across blocks. One plausible explanation for this pattern of results could be that participants become better ‘form-predictors’ as the experiment progresses. This, however, suggests that the emergence of an M100 effect for lexical-semantic prediction might have been a learning effect.

I already noted that while syntactic relatedness and form relatedness go hand in hand, this is not the case for lexical-semantic relatedness and form relatedness. I also pointed out that even in highly predictive sentence contexts, there are usually multiple lexico-semantically plausible continuations, which may or may not be related to each other in form. This is likely to result in differences between the two types of predictive processing (syntactic vs. lexical-semantic) in everyday language comprehension, where interlocutors typically do not show each other pictures corresponding to the words that they are about to utter. This concern with the present task regarding its ecological validity may explain why the M100 effect only emerged midway through the experiment: As pointed out above, the brain might not typically engage in predictive processing for the type of form properties that the M100 response is sensitive to, simply because it is not very often possible or helpful in everyday communicative contexts. In the +Predictive conditions of the present experiment, generating form expectations arguably was both possible and helpful. Possible in the
sense that there was always a one-to-one mapping between lexical-semantic representation and form, and helpful in that the generation of form-predictions allows for the early detection of mismatches, which in turn arguably facilitates faster responses. In other words, it is possible that the neural circuitry for generating form-based predictions based on lexical-semantic information is in place, but not frequently employed. Presenting people repeatedly with a situation where generating form-based expectations is rewarded might then motivate the engagement of the network that allows the direct top-down mapping of lexical representation onto visual form. The present results are not the first to show this type of flexibility in language processing. As was discussed in the previous chapter, for example, effects of grammaticality often attenuate as a result of extensive exposure to ungrammatical sentences, and effects of form typicality seem to disappear when participants are repeatedly presented with atypical forms.

In later time-windows, we did see an effect of congruence across the entire experiment, where lower amplitudes were observed for nouns that matched specific predictions as compared to all other conditions. This pattern of results supports an interpretation whereby lexical representations of nouns corresponding to specific pictures were preactivated / primed before presentation, as a result of which less processing cost was incurred when this noun was actually presented.

No later effects were found for mismatching trials in non-predictive contexts. In the Introduction, I already pointed out that it was not entirely clear what pattern of
results should be expected for this comparison. One possibility is that no lexical preactivation occurs in these cases, or maybe a bunch of nouns are preactivated, but these just happen not to always include the word that is actually presented.

For later components, the mismatch trials in the predictive contexts showed the same amplitudes as either the match or mismatch trials in the non-predictive contexts, whereas matching nouns in predictive contexts showed lower amplitudes than the other conditions. This is different from the M100 peak results, which showed amplitude reduction for predicted nouns and an enhancement for nouns that violated a prediction, when compared to M100 peak amplitudes to nouns for which no prediction was generated. Possible interpretations of these differences in effect directionality between the M100 response and later components will be discussed in the next chapter.

7.5 Conclusion

The results from Experiment 4 demonstrate that the M100 response is not solely sensitive to form features associated with syntactic category, but can also be responsive to violations of form-expectations that are based on lexical-semantic predictions. However, from the fact that an M100 effect of lexical-semantic violations only emerged in the second half of the experiment, it does appear that a crucial difference between processing syntactic and lexical-semantic violations. I here suggest
that this possibly results from inherent differences in the form-category relationship for syntactic category as opposed to lexical-semantic field. Possibly then, the typical precedence of syntax-related effects over lexical-semantic effects is due to differences in form-predictability rather than providing evidence in favor of a serial, bottom-up approach to language processing where syntactic computations take place before lexical-semantic processing (Friederici, 2002).
Chapter 8 – Neural correlates of predictive language processing

8.1 Introduction

In the preceding chapters, I provided results from a series of MEG experiments showing that visual cortex is sensitive to seemingly high-level factors. I suggested that early visual responses to word category violations and lexical-semantic mismatches follow from top-down modulation of visual cortex, resulting in the preactivation of form-features that are associated with the predicted syntactic category or word. In this chapter, I will present preliminary data pertaining to the neural correlates of such preactivation of visual cortex.

8.1.1 Bottom-up explanations of early sensory congruence effects

Considering extant neurolinguistic evidence, it appears unlikely that syntactic computation and lexical analysis are performed in sensory cortices. Further, it seems implausible that the M100 component could benefit from top-down feedback from higher cortical areas that support linguistic computations within the time-span before the M100 response, as this would predict sensitivity to syntactic and lexical factors for
words in isolation, i.e., in the absence of contextual information, a hypothesis that is not supported by previous results regarding the stimulus-sensitivity of the M100 component (e.g., Tarkiainen et al., 1999).

Thus, given both the timing and localization of the M100 response, in addition to its lack of sensitivity to linguistic factors for words in isolation, a predictive account seems much more plausible.

In theory, however, the findings discussed in the previous chapters could reflect bottom-up integrative effects, where word form, word category, and context compatibility are all evaluated during the M100 response: There is no direct evidence in the data presented so far that disproves this. For example, the results of Experiments 1-3 allow for an interpretation whereby word category analysis takes place based on visual information pertaining to form-typicality, only after which an attempt is made to integrate words into prior syntactic context, all during the M100 response. The most important implication of the results of Experiment 3 was that a bottom-up interpretation is not necessary, but it did not show that only a top-down account is possible.

For Experiment 4, it could be that modulations in M100 amplitude reflect a process whereby lexical representations are accessed and then matched against conceptual representations of preceding pictures. In fact, it has previously been argued that lexical access takes place as early as ~150 ms after stimulus-onset (Hauk and Pulvermüller, 2004), and there exist findings from the literature pertaining to
perceptual learning that even primary visual cortex can support representations of complex shapes after training and under task demands that require rapid recognition (Li, Piech, and Gilbert, 2004; Sigman et al., 2005). In line with this, Dambacher et al. (2009) argued that early lexical-semantic effects might only be found if the system is “pushed.” Under a perceptual learning interpretation then, the emerging M100 effect might have resulted from a process whereby e.g., temporary representations are created of a direct mapping between the word form and the conceptual representation associated with the presented pictures. Or maybe a limited set of lexical representations (the ones corresponding to stimuli used in the experiment, to which participants received repeated exposure) is copied over to visual cortex. One might ask what the functional interpretation could be of the effects observed in the N400 time-window if the M100 response already supports lexical representations, but there exist findings from other cognitive domains suggesting that rapid detection of mismatches at the ‘gate’ of visual processing might be enabled at the cost of representational redundancy: It has been argued that the formation of complex representations in lower level visual areas during the course of an experiment may occur in the form of duplicating representations from more upstream visual areas (Li, Piech, and Gilbert, 2004). In sum, there exist possible alternative accounts of the M100 effects reported here that do not strictly rely on form-feature expectations.

A crucial prediction of an account whereby the M100 effect does result from a mismatch with visual feature predictions is that relevant form features are somehow
pre-activated before a word is presented. As pointed out in the General Introduction, there exists some convincing indirect evidence from language processing that such preactivation does occur. Several scholars have reported mismatch effects on words preceding a predicted word, for example on prenominal adjectives or determiners that are inconsistent with the gender or phonological properties of the expected noun (Wicha et al., 2003; 2004; Van Berkum et al., 2005; DeLong et al., 2005). However, these studies still investigate consequences of predictive processing and not neural correlates of the preactivation process itself. In other words, while the existence of prediction in language processing has been convincingly demonstrated, to my knowledge no prior findings speak to the process of preactivating predicted representations.

In this chapter, I present preliminary results of an exploratory endeavor aimed at investigating whether we might be able to trace the neural correlates of anticipatory linguistic processing, aimed at finding direct evidence for the processes that support preactivation of predicted form representations in visual cortex.

8.1.2 A possible design to study prestimulus activity

The role of prediction in (neuro)cognition has recently received a lot of attention, and there now exist several relatively comprehensive accounts of the sources and
underlying mechanisms of predictive processing (e.g., Bar, 2007, 2009; Enns and Lleras, 2008; Bubic et al., 2010).

However, as pointed out above, while the existence of anticipatory processing in language comprehension is intuitive and widely accepted (Federmeier, 2007; Pickering and Garrod, 2007; Altmann and Mirkovic, 2009) convincing direct evidence is sparse. One reason for this is purely practical in nature. Crucially, investigating anticipatory processing requires an experimental design where visual/auditory input before the onset of the target word is maximally identical between conditions. Otherwise, effects in the prestimulus time-window for highly predictive vs. non-predictive contexts are confounded by differences between external stimuli. Given that context is exactly the factor that is likely to trigger predictive processing in some cases (he spread the warm bread with ___ ) but not in others (I like ___ ), equating input across conditions is next to unfeasible. In spite of this seemingly insurmountable limitation to experimental designs, some questions regarding the neural correlates of predictive processing can be asked with the data obtained from Experiment 4. Importantly, the target stimuli and most of the preceding context were identical across predictive and non-predictive contexts. As exemplified in Figure 20, presentation was held constant across the +PRED (predictive) and –PRED (non-predictive) conditions for a duration of 1500 ms between the offset of the picture (presented for 900 ms) and the onset of the target noun. Although the pictures are clearly different visual inputs,
brain activity associated with picture processing should subside well before the onset of the noun phrase (West and Halcomb, 2002; Willems et al., 2008).

Figure 20 - Time-line stimulus presentation (Experiment 4)

The timeline of the presentation of trials in Experiment 4. In this and all graphs shown below, the green line marks predictive contexts (presentation of a specific object); the grey line indicates non-predictive contexts (presentation of a grocery bag or Noah’s Ark). Pictures are presented for a duration of 900 ms (distinct across contexts), followed by 1500 ms of identical input across contexts.

In what follows, I first discuss possible mechanisms of expectation-based processing, showing data drawn from Chapter 7 (Section 8.2). For the M100 component, the results support an account that incorporates preactivation of predicted representations in conjunction with lateral inhibition of irrelevant representations. In contrast, for
those processes that underlie later effects (200 -400 ms), preactivation does not seem to involve simultaneous suppression of irrelevant representations.

Then, in Section 8.3, I present preliminary findings from minimum norm estimates of brain activity during the pre-stimulus interval, with the aim to investigate predictive processing directly. The results are promising: As would be predicted by previous literature, the findings suggest the involvement of ventro-medial prefrontal cortex, temporal cortex, and two occipital regions in the generation of context-based expectations as well as the preactivation of predicted representations.

### 8.2 Expectations and violations thereof: Mechanisms

Most evidence for expectation-based processing comes from studies that show preactivation of brain regions responsible for processing relevant types of sensory information. For example, gustatory cortices are activated in anticipation of food items (Simmons et al., 2005), and somatosensory cortex is activated before somatosensory stimuli are presented (Carlsson et al., 2000).

While these findings are promising, they only show that excitement of certain brain areas correlates with the nature of a particular type of upcoming percept: If a person is in ‘eating mode,’ gustatory cortices are excited, if she is in ‘face mode,’ fusiform gyrus is probed (Summerfield et al., 2006). Crucially, these results do not speak to the preactivation of representations of specific percepts (e.g., Asian face, or
the taste of a cupcake) and/or the possible suppression of irrelevant ones (e.g., my freckled face; the burn of jalapeño peppers). To take this a bit closer to home: In reading, it is likely that attention is directed toward visual stimulation as opposed to auditory stimulation, with the corresponding neural consequences thereof, but this does not automatically entail that specific representations of upcoming visual word forms are preactivated.

At this point it may be useful to note that albeit an intuitive distinction, contrasting attentive and predictive processing is not always empirically straightforward. Spratling (2008), for instance, compares predictive coding (Friston, 2003; 2005), a model that describes how prior information enables conditional expectations that are made available in a top-down fashion to lower-level visual areas, and the biased competition model (e.g., Desimone and Duncan, 1995; Desimone, 1996), which describes top-down effects in visual attention tasks. He demonstrates that at least some versions of these models are computationally equivalent, suggesting that the neural computations underlying the processes of attention and prediction respectively might be similar. For example, both models predict that preactivation of representations is accompanied by lateral inhibition of irrelevant ones.

A crucial difference, then, between a biased competition model and the types of ‘global attention’ effects discussed above, is that the former is designed to explain how attention can facilitate processing of certain visual features over others, whereas
the latter describe cases where certain sensory cortices show a general, Pavlovian-style, state of excitement as a function of anticipation.

It has also been suggested that mechanisms underlying prediction effects might be essentially identical to those that enable repetition priming effects (Lau, 2009), which is intuitive when considering that preactivated representations should be maximally similar to representations derived from incoming stimuli in a bottom-up fashion. Responses to expected words might thus be comparable to responses to repeated words.

Lau (2009) cites Grill-Spector, Henson, and Martin (2006), who review models designed to explain repetition priming effects, which typically involve a reduction of activity to the repeated stimulus compared to first presentation. Crucially, a subset of these models predicts that repeated presentation affects firing properties of the neurons that encode the predicted representation (less activity; less processing time), but not the state of neurons that encode competing representations.

We would not be able to infer from amplitude differences associated with predictive processing whether preactivation of predicted representations is competitive or not: in both cases there should be enhanced activation of relevant brain areas for predictive vs. non-predictive contexts. One might expect competitive preactivation to be more costly than non-competitive preactivation, but, needless to say, these cannot be compared directly. However, as shown below, competitive and non-competitive models do predict different patterns of activation at the target word depending on
whether or not suppression of irrelevant neurons forms part of the mechanisms underlying expectation.

To evaluate these models, it is important to determine a ‘baseline’ activation level in relevant brain areas, since both classes of models would predict less activity for predicted words compared to words that violate a prediction. In Experiments 1-3, none of the conditions matched the proper criteria, as different target words were presented in each comparison. In Experiment 4, however, the same nouns that were used in predictive contexts were presented under non-predictive but otherwise equal conditions. Therefore, I will here use the results of Experiment 4 to evaluate non-competitive vs. competitive accounts to explain the process by which predicted representations are preactivated.

### 8.2.1 Preactivation plus lateral inhibition: a competition account

One class of models, which Grill-Spector et al. (1996) refer to as ‘sharpening models,’ is represented by the biased competition model of visual attention, which I mentioned above (e.g., Desimone and Duncan, 1995; Desimone 1996). Under this approach (which, as pointed out, is mechanistically comparable to predictive coding models), top-down excitatory signals from pre-frontal to visual cortex bias the processing of certain representations over others. This excitatory signal acts to enhance the activation level of neuronal populations that encode relevant representations and
simultaneously suppress, by way of lateral inhibition, activity in neurons that represent irrelevant representations.

Figure 21 schematically describes the predicted excitation levels of neuronal populations in each condition of Experiment 4, both before stimulus onset and during presentation of the target noun (compared to resting state). Purple lines represent changes in the activation level of neuronal populations that encode form features associated with the word “apple.” I hypothesize that slope changes are reflected as an increase in brain activity in relevant cortices, indexing a cost associated with changes in activation level of neuronal populations therein.
Figure 21 - Predictions of competition models (Experiment 4)

Predicted excitation levels, under a competition account, of neuronal populations in each condition of Experiment 4. (A) -Predictive/Match, (B) -Predictive/Mismatch, (C) +Predictive/Match, and (D) +Predictive/Mismatch. Purple lines represent the activation level (relative to resting state) of neuronal populations that encode form features associated with the word “apple,” both during a predictive processing stage and during processing of the target word (demarcated by the dotted vertical line). Grey lines represent activation levels associated with unpredicted representations in each condition. Slope changes index changes in activation level of neuronal populations associated with (predicted) representations. The far-right panel indicates, for each condition, the predicted cost associated with bottom-up processing of the target word.
For the -Predictive conditions (Figures 21A and 21B), no change in activation level is expected as a function of predictive processing before the onset of the target noun. During presentation of the target word, some change in activity is expected due to bottom-up processing of “apple” in both cases, i.e., no difference in amplitude between match and mismatch conditions is expected for non-predictive contexts. In contrast, for the +Predictive conditions (21C and 21D), a competition model predicts changes in activation before the onset of the noun in those cortices where the predicted representations lie, indexing attentive excitatory and inhibitory processing. This should trigger more brain activity in the pre-noun time-window for both +Predictive conditions compared to both –Predictive conditions (compare the predictive processing stages for Figures 21A/B and 21C/D). During visual processing of the target noun “apple,” no significant changes in activation are expected for the +Predictive/Match cases (21C), as relevant neurons have already been preactivated. As a consequence, we would expect less brain activity during the presentation of the noun for this condition as compared to either of the -Predictive conditions. In contrast, for the +Predictive/Mismatch condition (Figure 21D), we would predict that the (re-)activation of suppressed neurons triggers a change in activation, causing amplitudes that are greater than for either of the –Predictive conditions, and much greater than for the +Predictive/+Match conditions. I have here also assumed a simultaneous suppression of preactivated neurons to below resting-state. However, complete deactivation is not a necessary consequence of the mismatch detection. Regardless, we
would predict the reactivation of suppressed neurons to be more costly than the activation of unsuppressed neurons (the –Predictive cases).

8.2.2 Preactivation without lateral inhibition: a facilitation account

As mentioned above, some models of repetition priming effects assume that reductions in brain activity for repeated stimuli as compared to first presentation affect only neurons that encode the predicted representation (fatigue models and facilitation models; Grill-Spector et al., 2006). Figure 22 (compare to 21) describes the predicted excitation levels of neuronal populations in each condition of Experiment 4 under such a facilitation only account.
Figure 22 - Predictions of facilitation models (Experiment 4)

Predicted excitation levels, under a facilitation account, of neuronal populations in each condition of Experiment 4. (A) -Predictive/Match, (B) -Predictive/Mismatch, (C) +Predictive/Match, and (D) +Predictive/Mismatch. Purple lines represent the activation level (relative to resting state) of neuronal populations that encode form features associated with the word “apple,” both during a predictive processing stage and during processing of the target word (demarcated by the dotted vertical line). Grey lines represent activation levels associated with unpredicted representations in each condition. Slope changes index changes in activation level of neuronal populations associated with (predicted) representations. The far-right panel indicates, for each condition, the predicted cost associated with bottom-up processing of the target word.
In the absence of suppression of irrelevant representations in the prestimulus interval (grey line in Figure 22C, compare to 21C; purple line in Figure 22D, compare to 21D) we should still see higher amplitudes for the +Predictive cases compared to the -Predictive cases in the prestimulus interval, albeit by a smaller magnitude (assuming that competition is more costly than mere priming). Such magnitude differences cannot be evaluated with the present data set however.

Similarly to a biased competition model, facilitation models would predict no differences between the match and mismatch conditions in the absence of prediction, as both cases merely involve the activation of relevant neuronal populations during the input (compare Figures 22A/B and 21A/B). Both models would also predict that activity at the target noun in +Predictive/Match cases is reduced compared to either of the -Predictive conditions (compare Figures 22C and 21C), as the neurons that represent the predicted representation have already been preactivated.

However, in a model without lateral inhibition, encountering a word that does not match the predicted representation does not necessitate the activation of neurons from a state below resting level, as activity in these neurons is not assumed to have been suppressed before stimulus onset. Rather, processing violations of predictions might simply involve the activation of neurons that encode the incoming presentation, similar to processing words for which no prediction was generated at all (compare Figures 22A/B and 21D). Presumably, this process would be less costly than engaging previously suppressed neurons (compare Figures 24D and 23D). As a result, while a
biased competition model would predict more brain activity for violations of predictions compared to words for which no predicted representation was preactivated, a facilitation model without lateral inhibition only predicts reduction for words that satisfy an expectation (Figure 22C) compared to all other conditions.

Figure 23 summarizes the predictions of competition and facilitation models respectively regarding the relative amplitudes of brain activity associated with processing target words in each condition. For purposes of visualization, connecting lines are drawn between the match conditions and mismatch conditions respectively.
PREDICTIONS FOR ACTIVITY AT TARGET WORD

a. preactivation & suppression

b. preactivation w/o suppression

Figure 23 - Prediction graphs for target word amplitudes (Experiment 4)
Relative activity amplitudes for each condition (grey: -Predictive; green: +Predictive; blue: match; red: mismatch) as predicted by competition models (a) and facilitation models (b). Connecting lines between amplitudes are given to illustrate the relationship between match and mismatch conditions respectively between comparisons.

As said above, a model that assumes both preactivation of predicted representations and suppression of competing representations would predict (a) no differences between the match and mismatch conditions for non-predictive contexts; (b) reduced activity for cases where predictions are satisfied; and (c) increased activity compared to all other conditions for words that violate a prediction. In contrast, a model that
does not incorporate lateral inhibition would predict reduced activity for anticipated words compared to all other conditions, and no differences between words that violate predictions compared to words in non-predictive contexts.

While the models presented above are derived from the attention and predictive coding literature on the one hand, and repetition priming models on the other, they are also compatible with models of lexical activation. Lexical selection has been found to be competitive between phonologically related words (e.g., Pylkkänen, Stringfellow, and Marantz, 2002): At the presentation of a word, a cohort of phonological neighbors is activated together. Then, once one activated representation ‘wins’ for lexical selection, all other candidates are suppressed, possibly even below resting level. It has been shown that phonological competition effects in lexical selection do not surface before ~350 ms after the onset of a word. However, in the context of prediction we would expect competition effects between phonologically related words to take place before the onset of the stimulus, namely during the preactivation of the predicted representation.

In contrast to phonological neighbors, semantic relatives do not compete for lexical access, rather facilitating each others’ recognition, possibly by way of spreading activation through the lexicon (e.g., Collins and Loftus, 1975).

This means that preactivation of predicted representations should be competitive between representations for phonologically related representations and non-competitive between semantically related representations, following the pattern of
changes in activation grossly similar to the ones presented in Figures 21 and 22 respectively.

I have suggested that the M100 effect indexes a mismatch at the form feature level, while later effects reflect ‘true’ lexical access. While both phonological and semantic factors affect lexical selection, which has been argued to occur in the N400 time-range (Pylkkänen and Marantz, 2003), it is not unlikely that phonological effects are driven by more low-level form-related processes. For experiment 4, then, we might expect that those predictive processes that operate on form representations (the M100 component) are competitive in nature, while predictive processes that operate on semantic representations (arguably the later components) are facilitatory in nature.

8.2.3 M100 and left-anterior responses to target words

Figure 26 presents the peak amplitudes over M100 sensors (24A) and left-anterior sensors in the 200-400 ms time-window (24B). Only trials presented in Block 2 are included here (see Chapter 7 for the full design).
RESULTS FOR ACTIVITY AT TARGET WORD

a. M100 response preactivation & suppression
b. LA sensors (200-400ms) preactivation w/o suppression

Figure 24 – Competition vs. facilitation: M100 and left-anterior sensors (Experiment 4)
Relative amplitudes for each condition (grey: -Predictive; green: +Predictive; blue: match; red: mismatch) as observed in peak activity for the M100 sensors (a) and left-anterior sensors in the 200 – 400 ms time window (b). Connecting lines between amplitudes are given to illustrate the relationship between match and mismatch conditions respectively for each comparison.

As pointed out above, neither a competition account nor a pure facilitation account would predict differences in amplitude for the -Predictive comparison. Indeed, we see that similar amplitudes for matching and non-matching words in the -Predictive manipulations are observed for both the M100 response and left-anterior sensors.
The relationship between amplitudes to the +Predictive/Mismatch condition and those to the -Predictive conditions, however, does differ between the M100 response and the later component.

M100 amplitudes (Figure 24A) appear to reflect a decrease of sensory activity associated with the processing of predicted words and an increase of visual processing in response to words that violate a prediction, supporting an account where processing unpredicted words in predictive contexts involves re-engaging previously suppressed neurons.

In contrast, left-anterior sensors in the 200-400ms time-window did not exhibit increased amplitudes for unpredicted (+Predictive/mismatch) words compared to words presented in non-predictive contexts. Only reduced activity for predicted words compared to all other contexts was observed, supporting a general facilitation account without prestimulus competition between representations.

These findings suggest that the mechanisms underlying the preactivation of visual features are qualitatively different from those that subserve semantic activation, the latter being facilitory only.

As the primary goal of this dissertation is to explain early visual effects of expectation violations, in what follows I focus on the M100 response. Figure 25 exemplifies a simplified version of the processes that are assumed to affect M100 amplitude for the +Predictive/Mismatch condition under a competition account. For illustrative purposes, ‘form features’ are here shown as each neuron representing a
letter. Recall, however, that it remains unclear which level of word form representation is supported by the M100 generator. During presentation of the prime picture (left model brain), the lexical entry for \{banana\} is preactivated, which triggers top-down modulation to visual cortex. This causes excitation of relevant letters (blue bubbles; \(a, b, n\)) and suppression of competitors (e.g., \(p, l\)). Then, during presentation of the unexpected target word “apple” (right model brain), visual analysis triggers the activation of \(p\) and \(l\) (possibly in addition to a suppression of \(b\) and \(n\)).
PREACTIVATION AND SUPPRESSION: VIOLATIONS OF PREDICTIONS

(A) Predicted excitation levels, under a competition account, of neuronal populations in the M100 generator for the +Predictive/Mismatch condition of Experiment 4. The blue line represent the activation level (relative to resting state) of neuronal populations that encode form features associated with the word “banana,” both during a predictive processing stage and during visual processing of the target word (demarcated by the dotted vertical line). The red line represents activation levels associated with the neuronal populations that encode “apple.” Slope changes index changes in activation level of neuronal populations associated with (predicted) representations. (B) Cartoon brain model representing the underlying mechanisms that might explain expectation and violation effects at the M100 response.
In sum, we expect (a) increased visual activity for “apple” in the +Predictive/Mismatch and reduced amplitudes on the target noun for the +Predictive/Match condition; and (b) enhanced activity in visual cortex in pre-stimulus activity for +Predictive conditions. (a) was confirmed by the results above (Figure 24A); (b) will be tested below. There already exists some direct evidence pertaining to preactivation of representations in visual cortex: Stokes and colleagues find that visual imagery activates the same neural representations that are active in response to actual visual presentation of an imagined shape (Stokes, Thompson, Nobre, and Duncan, 2009).

8.3 Tracing preactivation of visual cortex in predictive language processing

8.3.1 Analysis: minimum norm estimates

In all experiments discussed above, I analyzed amplitude variations of the visual M100 response to words that either violated or met predictions. Activity at the M100 component was either determined by comparing dipole waveforms (Experiments 1-3) or by investigating activity recorded at representative sensors (Experiment 4). Neither of these methods, however, is particularly appropriate to address questions regarding activity in visual cortex for any other time-window than during the M100 response to
an incoming word, mainly because brain activity does not exhibit discrete components in the time-window immediately preceding stimulus onset. As a result, no reliable dipole modeling or sensor selection can be performed. One might propose to investigate sensory activity before word presentation by extrapolating the dipole models or sensors representing the M100 response to the critical word. However, M100 dipole models characterize activity during a specific time-window; a high goodness-of-fit is a crucial prerequisite to justify interpreting any variation in waveform amplitude. And it is not guaranteed that activity at sensors that are representative of the M100 field pattern reflects activity generated in visual cortex in any other time-window than during the M100 response to the critical word. To illustrate, some of the sensors at the center of the re-entering field of the M100 component are also representative of later components, as can be seen in Figure 15.

A reasonable alternative approach to investigating prestimulus brain activity is to use minimum norm estimates (MNEs) of neural activity across the whole brain without posing temporal or spatial restrictions. For the data presented below, MNEs were calculated in BESA 5.1 based on activity from 1426 regional sources evenly distributed in two shells 10% and 30% below a smoothed standard brain surface. Regional sources can be seen as two orthogonally oriented dipoles in the same location; total activity was thus calculated as the root mean square of the source activities. The minimum norm images were depth-weighted as well as spatio-temporally weighted.
In the next few sections, I show whole-brain minimum norm estimates calculated for predictive vs. non-predictive contexts in the time-window preceding the onset of the target noun, and match vs. mismatch trials during the presentation of the noun. First, I focus on the question of whether visual cortex is preactivated as a function of prediction, as would be expected under the Sensory Hypothesis. Following this, I present data suggesting that both left-temporal lobe and medial prefrontal cortex are recruited in predictive processing.

### 8.3.2 Results

In all figures below, colored regions indicate brain areas where differences in brain activity for predictive vs. non-predictive context were reliable (p < .05) for a duration of least 50 ms, and were observed in at least 10 adjacent sources (red: more activity for the +Predictive conditions as compared to the -Predictive conditions; blue: less activity for the +Predictive conditions). First, I show minimum norm estimates for data filtered between 1 and 40 Hz, corresponding to the filters that were used in all experiments presented above when investigating brain activity during the target word. Then, I show results for data filtered at different frequency bands, following previous research that suggests that prediction selectively affects activity in lower frequency bands.
a. Brain activity filtered between 1-40 Hz

Figure 26 displays an occipital view of the results of a whole-brain analysis over minimum norm estimates (MNEs) comparing +Predictive to –Predictive contexts from the onset of the determiner preceding the target noun to the onset of the target noun, divided into 150 ms intervals. Only trials from Block 2, where predictive and non-predictive contexts were intermixed, are included, given that this is the Block where an M100 effect of congruence was observed (see Chapter 7).

![Figure 26 - Prestimulus activity; whole-brain minimum norm estimates (Experiment 4)](image)

*Results of a whole-brain MNE analysis for data filtered at 1 – 40 Hz. Colored regions indicate brain areas where differences in brain activity for predictive vs. non-predictive contexts were reliable (p < .05) for a duration of least 50 ms, and observed in at least 10 adjacent sources (red: more activity for the +Predictive conditions compared to the -Predictive conditions; blue: less activity for the +Predictive conditions).*
As can be seen, under these thresholds no differences are found between predictive and non-predictive contexts. In the next section, I explore the possibility that prediction effects might not surface when looking at data across different frequency bands.

b. **Brain activity concentrated in different frequency bands**

Lau (2009) argues that predictive processing is probably not time-locked in the same way as e.g., early visual responses are. As a result, there might be a lot of temporal inconsistency between trials as well as participants with respect to the onset of predictive effects. Thus, such effects would not survive the usual averaging procedures. These concerns are especially relevant for changes in brain activity that take place in mid- to high- frequency ranges, which may be relevant ranges to predictive processing. For example, Gregoriou et al. (2009) found that attention in the monkey brain affects spiking patterns in the 40-60 Hz range both in the frontal eye-field and visual cortex (recall that attentive processing has been linked to predictive processing [e.g., Spratling, 2008; see also above]). Such high-frequency modulations are likely to be obscured by averaging even for processes that are time-locked to e.g., visual stimulation. Interestingly, however, the changes in spiking patterns as a function of attention resulted in a sustained increase in activation. It has been shown that long-lasting spiking changes in higher frequency ranges of the type observed by
Gregoriou et al. (2009) may surface as amplitude differences at lower frequencies, capturing the envelope of high frequency modulations (as opposed to their internal structure). In fact, Kirmizi-Alsan et al. (2006) have shown that continuous attention affects modulation in very low frequency ranges (below 4 Hz). Such long-lasting effects should be able to survive some inconsistencies in onset time. But, given that low frequency ranges are included in the minimum norm estimates reported in Figure 26, then shouldn’t we have found some amplitude differences?

It may be the case that sustained effects are obscured by temporally jittered or non-event related changes in activity at intermediate frequencies (~40 to 10 Hz). As an initial investigation into this option, I performed analyses over activity extracted from different low-frequency bands: delta (1-3 Hz), theta (4-7 Hz), and beta (8-12 Hz). Of particular interest in this respect is a study by Von Stein, Chlang, and Konig (2000), who suggest that processes of top-down modulation are likely to surface in the theta frequency range, whereas local computations have been associated with higher frequency (gamma) activity. Thus, we might expect preactivation effects to surface when restricting analysis to 4-7 Hz specifically.

Results are presented in Figure 27. Panel 1 shows whole-brain minimum norm estimates for prestimulus activity comparing predictive and non-predictive contexts in each frequency band. Panel 2 displays post-stimulus whole-brain minimum norm estimates of changes in brain activity for the match vs. mismatch comparison in the +Prediction conditions (Block 2). Colored regions indicate brain areas where
differences in brain activity for each comparison were reliable (p < .05) for a duration of least 50 ms in the +Predictive vs. –Predictive comparison (before stimulus onset), and 25 ms for the Match vs. Mismatch comparison (during presentation of the noun), and were observed in at least 10 adjacent sources (red: more activity for the +Predictive conditions compared to the -Predictive conditions; blue: less activity for the +Predictive conditions).
Figure 27 - Whole-brain minimum norm estimates for prestimulus and target word brain activity at different frequency ranges (Experiment 4)

Results of the whole-brain analysis over MNEs for data filtered at 13-30 Hz (Beta range; 27A), 8-12 Hz (Alpha range; 27B), 4-7 Hz (Theta range; 27C), and 1 – 3 Hz (Delta range; 27D) respectively. Colored regions indicate brain areas where differences in brain activity were reliable (p < .05). Panel 1. For a duration of least 50 ms, observed in at least 10 adjacent sources, Red: more activity for the +Predictive conditions than the -Predictive conditions; Blue: less activity for +Predictive conditions. Panel 2. Duration: + 25 ms, Adjacent sources: 10, Red: more activity for the Mismatch condition than Match condition; Blue: less activity for the Mismatch conditions.
As can be seen from Figure 27, differences in brain activity before stimulus onset as a function of context are only found in the lower frequency ranges, namely theta (Figure 27.3.A), and delta (Figure 27.4.A). No effects of context are observed under these thresholds in the alpha or beta ranges (Figures 27.1.A and 27.2.B respectively). In the theta-range, we see enhanced activity for predictive contexts as compared to non-predictive contexts in visual cortex right before the onset of the noun. In the delta range, in contrast, we see long-lasting enhanced activation of visual regions in non-predictive contexts. As can be seen in Panel 2, however, effects of congruency in visual areas during the presentation of the noun are found only in the theta-band, and not in the delta-band (notice that visual effect during the target word seems to spread from a posterior region before 150ms to a slightly more (left) lateral area after 150ms). This leads us to suspect that, rather than reflecting the preactivation of predicted representations, the enhanced activity for –predictive contexts in the delta-band might index an increased level of uncertainty about the upcoming stimulus (potentially an attention effect; Kirmizi-Alsan et al., 2006), but this interpretation is strictly speculative.

In sum, the data shown in Figure 27 suggest that brain activity concentrated in the theta-range in particular is sensitive to prediction: we find visual effects in this frequency band both before the onset of the noun as a function of context, and during the presentation of the noun as a function of whether or not a prediction is satisfied.
8.4  Predictive effects beyond visual cortex: ventro-medial
prefrontal cortex and the left temporal lobe

In addition to preactivation of visual cortex as a function of expectation, most models would predict that ‘higher’ cortical regions are involved in predictive processing. Although there exists no direct evidence pertaining to where contextual predictions in language processing might be generated, there are several brain areas that one might expect to be recruited. In this section, I discuss which brain regions beyond and before visual cortex might be involved in anticipatory linguistic processing.

8.4.1 Possible neural sources of predictive processing

In this section, I discuss previous findings regarding both the sources of predictive processes and the preactivation of higher-level representations.

a. Prefrontal cortex and prefrontal-occipital connectivity

Various regions of prefrontal cortex can be identified as likely sources of anticipatory processing (e.g., Bar, 2007; Bubic et al., 2010). Bar (2007) specifically proposes the involvement of medial prefrontal cortex (mPFC) in generating predictions, an area that
has been associated with cognitive control broadly, and evaluating potential outcomes in particular (Ridderinkhof et al., 2004). Ventral regions of the medial prefrontal cortex have been implicated in generating anticipatory and top-down processing for visual object recognition (e.g., Summerfield et al., 2006; Bar et al., 2006). Further, activity in anterior cingulate cortex (ACC) has been found to correlate with preparatory processing for upcoming perceptual events (Gómez, Vaquero, and Vázquez-Marrufo, 2004) as well as the likelihood of prediction error (Brown and Braver, 2005).

Prefrontal regions have further been implicated in top-down modulation of visual cortex as a function of attention (e.g., Desimone and Duncan, 1995; Lamme, 2003; Gregoriou, Gotts, Zhou, and Desimone, 2009), and it has been suggested that enhanced phase synchronization of neuronal populations in prefrontal regions predict visual processing as well as behavioral responses (e.g., Engel, Fries, and Singer, 2001; Liang et al., 2002). In sum, we might expect enhanced activity in prefrontal areas as a function of predictive processing.

A number of studies have demonstrated top-down as well as predictive effects on visual processing, and several scholars have proposed theoretical accounts of how these processes might be implemented (e.g., Albright and Stoner, 2002; Rao and Ballard, 1999; Lee and Mumford, 2003; Bueti, Bahrami, Walsh, and Rees, 2010; Alink et al., 2010; Stokes, Thompson, Cusack, and Duncan, 2009).
Further, a few studies have investigated the connection between prefrontal areas and sensory regions in anticipatory processing. For example, Summerfield and Koechlin (2008) compared predictive and non-predictive contexts in a low-level visual manipulation and found increased backward connectivity from Fusiform Gyrus (FG) to primary visual cortex (V1) in predictive contexts only (compared to non-predictive contexts). They further report increased feedforward connectivity between V1 and FG for prediction mismatches, suggesting an interaction between top-down and bottom-up processing streams. As none of the experiments discussed in the previous chapters have found effects in FG, we here expect no predictive effects in FG.

Another study investigated contextual prediction in resolving perceptual ambiguity in object identification and found an increase in top-down connectivity between prefrontal regions and occipito-temporal cortex depending on context (Summerfield et al., 2006). However, in this research “context” was defined very globally (in terms of experimental blocks) and was strictly induced by the experimental task. In natural language processing, context is dynamic and local. Also showing a relationship between higher cortical regions and visual cortex, the study by Gregoriou et al. (2009; discussed above) found that sustained changes in high-frequency spiking patterns in visual cortex were immediately preceded (by ~ 50 ms) by similar changes in the frontal eye field (FEF), suggesting a relationship between these two areas.
b. Left temporal cortex and inferior frontal gyrus

Summerfield and Koechlin (2008), and Summerfield et al. (2006), found effects of visual object prediction in Fusifurum Gyrus, arguably reflecting the preactivation of visual object representations. In the case of linguistic prediction then, we would equally expect effects in regions that support linguistic representations. For lexical-semantic prediction, these areas could include the middle temporal lobe (mTL) and anterior and posterior inferior frontal gyrus (IFG) specifically, regions that have been implicated in lexical access, retrieval, and selection respectively (Lau, Phillips, and Poeppel, 2008). For example, based on findings from e.g., Thompson-Schill et al. (1997) and Baumgaertner, Weiller, and Büchel (2001), Lau (2009) proposes that anterior IFG might mediate semantic retrieval by way of top-down modulation to (middle-posterior parts of) left temporal cortex, where lexical representations are arguably housed.
8.4.2 Results

Figure 28 displays whole-brain minimum norm estimates for both the +/-Prediction comparison (Panel 1) and the Match/Mismatch comparison (Panel 2; +Predictive/Block 2 conditions only), exposing the left hemisphere and prefrontal cortex collapsed across 150 ms time-intervals. Red patches on the surface of the model brains indicate regions that were more active for predictive context than non-predictive context (prestimulus window; Panel 1) or more activity for the mismatch than the match trials (during presentation of the noun), for at least 40 consecutive milliseconds, and in a minimum of 10 neighboring sources.
Figure 28 - Whole-brain minimum norm estimates for prestimulus and target word brain activity in theta band (Experiment 4)

Results of whole-brain analysis over MNEs for data filtered at 4-7 Hz (Theta range). Colored regions indicate brain areas where differences in brain activity were reliable (p < .05), for a duration of at least 40 ms, observed in at least 10 adjacent sources. Panel 1. Red: more activity for the +Predictive conditions than the -Predictive conditions; Blue: less activity for +Predictive conditions. Panel 2. Red: more activity for the Mismatch condition than Match condition; Blue: less activity for the Mismatch conditions.

The following pattern emerges from Figure 28. First, in the prestimulus time-window (Panel 1) we see that two sources are more active for the predictive conditions between 300 and 150 ms before the onset of the noun: the left middle temporal lobe (mTL) and a small portion of the ventro-medial prefrontal cortex (vmPFC). As pointed out in the previous section, left mTL has been proposed to support lexical representations (see e.g., Lau et al., 2008 for review). Thus, this difference in
activation potentially indexes the cost associated with preactivating the lexical representation of a predicted word. As discussed above, vmPFC is one of the regions in prefrontal cortex that has been previously associated with predictive processing, as well as visual attention.

mTL and vmPFC also respond to whether or not a given noun matches the prediction (Panel 2), although the mismatch response is much greater and more broadly distributed than the prediction effect. Increased activity for mismatch trials was also found in e.g., left IFG, a region that has been implicated in numerous linguistic processes, including word generation (Thompson-Schill et al., 1997), working memory, and even specifically prediction for lexical representations, as discussed above (Lau, Phillips, and Poeppel, 2008).

This pattern of activation is exactly what would be expected, namely that violations of expectations as a function of predictive processing are observed in the same brain areas that are active before the stimulus comes on. For example, if lexical-semantic predictive processing involves the preactivation of lexical representations in the left temporal lobe, a prediction mismatch should affect lexical processing during the presentation of the lexical item that forms a mismatch. Similarly, the same brain regions in ventro-medial prefrontal cortex that have been associated with prediction and visual attention (before stimulus-onset) have been shown to be sensitive to prediction error.
Thus, Figure 28 and Figure 27 together suggests that enhanced activation in left temporal and prefrontal cortex for predictive contexts is followed by enhanced activity in visual cortex. As for amplitude differences at the target word as a function of congruence, we see a reverse pattern: Peak activity in the VC region is significantly increased for mismatch trials as compared to match trials between 100 and 200 ms after the onset of the noun (Figure 27). Following this effect, we see increased activity for the mismatch condition in both mTL and vmPFC (Figure 28) between ~350 and 600 ms.

In sum, we see that the same brain regions that exhibited enhanced activation for predictive contexts before the onset of the noun show effects of congruence during the target word, but in reverse order. This is exactly the pattern that would be expected if these are indeed effects resulting from predictive processing: predictive processes would be expected to feed back from higher to lower cortical regions, whereas prediction error, or mismatch, responses should follow a bottom-up processing stream from low-level to higher-level representations (Friston, 2005; Summerfield and Koechlin, 2008).
8.5 **Discussion**

The data presented in this Chapter implicate a number of regions in predictive processing. In the next few sections, I discuss these in turn, in terms of possible functional interpretations.

8.5.1 **Preactivation of predicted representations: visual cortex and left temporal cortex**

In all results presented above, we saw effects of prediction in the left middle temporal lobe (mTL), followed by a prediction effect in visual cortex. The reverse pattern (VC effect first, then left temporal effect) was observed for effects of congruence. As pointed out above, the left mTL has been proposed to support lexical access (see e.g., Pylkkänen & Marantz, 2003; Lau et al., 2008 for review). Thus, the enhanced activation of the left mTL before stimulus-onset may well reflects preactivation of the predicted lexical representation, based on which form feature estimates are generated. In other words, our findings suggest preactivation as a result of lexical-semantic prediction at both the level of lexical representation (mTL) and word form features (VC).

In sum, this set of findings provides relatively direct evidence in support of the Sensory Hypothesis: the M100 effects observed in the preceding chapters may indeed be a result of violations of form-based expectations.
8.5.2 Ventro-medial prefrontal cortex

A number of medial prefrontal regions were mentioned as plausible candidates to be involved in predictive processing. We here found activity associated with anticipatory processing in the ventral part of mPFC in particular. A range of non-language specific functions have been attributed to mPFC, including, as mentioned above, attention (Desimone and Duncan, 1995) predictive processing (Bar, 2007), and top-down effects in object perception (Bar et al., 2006). Miller and Cohen (2001) suggest that prefrontal cortex serves an integrative function between a number of cortical regions (Miller and Cohen, 2001), and Fuster (2001) has proposed that temporal alignment is one of the crucial functions of prefrontal cortex. In view of all of this, it is not unlikely that vmPFC would be involved in communicating expectations to occipital regions. With respect to congruence effects at the target noun, we also saw that vmPFC showed mismatch effects, which would be expected under a prediction account.

As vmPFC has not been previously implicated in lexical access, it may rather be the case that vmPFC serves as a vehicle for the translation of lexical-semantic representations into form-based estimates. Many of the numerous functions that have been attributed to vmPFC, ranging from ‘abstract’ cognitive functions such as theory of mind to multi-sensory integration, are compatible with an interpretation where a primary role of vmPFC is to support such a binding function between a number of regions and levels of representations (see also Wood and Grafman, 2003, for a review.
of evidence in favor of a view where vmPFC performs a linking function between representational levels). However, this interpretation is speculative at best when regarding the present data, given that the effects in vmPFC and mTL occurred roughly around the same time.
Chapter 9 - Overall summary and conclusions

I started out by asking how language processing can be so fast. Or, more specifically:
How is it possible that some effects of syntactic incompatibility occur as early as
100ms after word-onset?

To begin investigating this puzzling phenomenon, I formulated the Sensory
Hypothesis (Dikker, Rabagliati, and Pylkkänen, 2009), which guided the research
presented in this dissertation. This hypothesis consists of three predictions:

1. The parser is ‘eager’ and makes continuous word category predictions.
2. These predictions include form-estimates associated with a predicted syntactic
category.
3. If these predictions are violated, this causes a mismatch effect in sensory
cortex.

The first prediction of this hypothesis was already confirmed by previous findings
showing that anticipatory processing plays a crucial role in language comprehension
(e.g., Altmann and Kamide, 1999; 2009; DeLong et al., 2005; Van Berkum et al.,
2005; Wicha et al., 2003; 2004; Lau et al., 2006; Staub and Clifton, 2006).
In Experiment 1, we found evidence in support of the third prediction of the Sensory Hypothesis: At least some effects of syntactic incompatibility appear to be generated in visual cortex (the visual M100 response in MEG). This experiment was the first to demonstrate that early visual responses to word forms can be influenced by prior syntactic context.

Clearly though, such a finding in and of itself does not explain why syntactic effects take place so early. In fact, this result generated a whole series of new questions: What types of representations are supported by the M100 generator? Is this an effect resulting from contextually generated predictions? And if so, can we trace where these predictions might be generated?

The bulk of the research presented here was dedicated to fleshing out the details of prediction 2 of the Sensory Hypothesis. First, in chapters 3-6, I discussed results from three experiments whose goal it was to assess the nature of the word category cues available to the M100 generator. Stated in terms of the Sensory Hypothesis: Does an M100 effect arise because of a prediction mismatch in form representations? Or does visual cortex support operations on ‘higher’ representational levels, such as word category analysis or lexical access?

In Experiment 1, an M100 effect was only obtained for word category violations where the target word contained a closed-class category marking morpheme. However, words that did not generate early visual effects in unexpected context were category ambiguous (bare stems). Experiment 2 addressed whether word
category violations involving category-unambiguous nouns (like *tree*, as opposed to *report*) without an overt closed-class category marking morpheme (*tree* as opposed to *report-ed*) generate an M100 effect. If so, this would indicate that some form of lexical access takes place at the M100 component, since ambiguity resolution is arguably performed at the lexical level. Experiment 2 found no M100 effect for category-unambiguous nouns like *tree*, suggesting that lexical access does not take place during the M100 response.

A follow-up question, then, was if an M100 effect might be exclusive to words containing closes-class morphemes, since they are frequent, short, relatively salient, and they form only a small subset of the lexicon. Thus, it may be conceivable that the visual system of experienced readers might be particularly tuned to detect them. However, if CCMs were special at this stage of analysis, then it would follow from the pattern of results obtained in Experiments 1 and 2 that the M100 effect reflects a true word category mismatch, and not a mismatch with form predictions. In other words, the M100 generator would have to support word category analysis based on the presence of category marking closed-class morphology.

Experiment 3 found that the M100 response is sensitive not just to closed-class morphemes, but also to form features that are probabilistically associated with word category (Dikker, Rabagliati, Farmer, and Pylkkänen, 2010). In view of Prediction 2 of the Sensory Hypothesis, this finding is important not only because it informs us about the types of word category indices to which the M100 response is sensitive, but
also that a bottom-up account of early visual effects of syntactic incompatibility where word category analysis takes place during the M100 response, is no longer warranted.

This then led into the question why early sensory effects have not been previously obtained for lexical-semantic violations. Might sensory cortices be exclusively tuned to detect syntactically relevant form features? Such a finding could either indicate that there exists a qualitative difference between syntactic and lexical-semantic processing during the first stages of visual analysis, or that there exist qualitative differences between syntactic and lexical-semantic predictive processing: potentially, only syntactic predictions are translated into form-expectations.

In other words, I asked whether the Sensory Hypothesis applies only to syntactic predictions, or to linguistic predictions in general, testing an alternative to the formulation of Prediction 2 of the Sensory Hypothesis:

2’ These predictions include form-estimates associated with the predicted linguistic representation.

Experiment 4 (Chapter 7) investigated whether the M100 generator discriminates between word form properties associated with syntax as opposed to lexical-semantics. I suggested that violations of lexical-semantic expectations do not typically result in sensory effects because lexical-semantics and word form do not usually go hand in
hand, while syntactic category and word form do (e.g., by way of the form typicality measure used in Experiment 3).

This hypothesis was tentatively supported by the finding that an M100 effect can indeed be obtained for violations of lexical-semantic predictions, but only when a word violates a prediction for a single lexical item. However, an M100 effect for lexical-semantic mismatches emerged only during the second half of the experiment, tentatively supporting the claim that comprehenders may not typically engage in generating form-estimates based on lexical-semantic predictions because of the lack of form correlates of these predictions in the majority of cases. Regardless, the results of this study did suggest that form-based predictions are not generated solely on the basis of syntactic category predictions, tentatively confirming Prediction 2’, and warranting a broadening of Prediction 1 (in fact it was too narrow to begin with, given that a number of previous findings have suggested a crucial role of prediction in lexical-semantic processing):

1’ The parser is ‘eager’ and makes continuous predictions, at various levels of representation.

2’ **When possible**, these predictions include form-estimates associated with the predicted linguistic representation.
However, while an explanation of the findings from Experiments 1-4 in terms of prediction is clearly the most plausible one, the results in and of themselves do not prove that predictive processing underlies early sensory effects. This can be done only by way of a direct demonstration of the existence of predictive processing that enables the generation of form-estimates.

Chapter 8 presented an exploratory investigation into the neural correlates of predictive language processing, asking which brain regions might show differences in amplitudes depending on whether context is predictive, during the time-interval preceding (un)expected nouns (using the data obtained from Experiment 4). Preliminary results suggest that the generation of predicted form representations associated with lexical-semantic predictions recruits ventro-medial prefrontal cortex, left middle temporal cortex and (the vicinity of primary) visual cortex. Importantly, the same brain areas were more active for words that violated lexical-semantic expectations than for words that satisfied predictions, but in reverse order.

In the context of the Sensory Hypothesis, the most important finding is that a posterior occipital region was more active for predictive contexts immediately preceding the onset of the target noun, and that this same region showed effects of congruence at the target noun in the same time-window as the M100 response, as well as the same pattern of amplitude modulations. This is highly suggestive that this region supports predicted representations of the type of form features to which the M100 component is responsive. In addition, middle temporal cortex showed increased
activation for predictive contexts before the effect in visual cortex, potentially indexing the preactivation of lexical representations from which these form features are derived.

In Chapter 8, I also investigated the mechanisms underlying predictive processing. The pattern of results observed in Experiment 4 suggest that the preactivation of predicted form representations in visual cortex involves a combination of excitation of target form representations and a simultaneous suppression of irrelevant form representations, consistent with attention models of biased competition (e.g., Desimone and Duncan, 1995) and predictive coding models (e.g., Friston, 2005). In contrast, the preactivation of lexical-semantic representations (arguably housed in left temporal regions), appears to involve facilitation only (i.e., without lateral inhibition of unpredicted representations.

In summary, the research presented in this dissertation allows us to fill in some of the blanks of the Sensory Hypothesis:

1. The parser is ‘eager’ and makes continuous linguistic predictions, at various levels of representation.
   - Predictive processing recruits ventro-medial prefrontal regions as well as brain areas that operate on the level of analysis for which predictions are generated.
2. When possible, these predictions include form-estimates associated with the predicted linguistic representation.

- Pre-activation of form representations results in the activation of relevant representations and simultaneous suppression of irrelevant ones. This process is reflected by increased brain activity in sensory cortex prior to sensory stimulation.

3. If these predictions are violated, this causes a mismatch effect in sensory cortex, followed by mismatch effects in cortical areas that are responsible for analyzing ‘higher’ level predicted linguistic representations (such as syntactic category or lexical-semantic content).

These hypotheses might provide us with some tools to replace the question marks in Figure 1 in the Introduction, repeated here.

![Figure 29 - Research questions: predictive language processing in the brain](image)
Figure 30 presents a cartoon model of predictive processing (A) and prediction error responses (B), summarizing the findings and the proposed mechanisms discussed in previous chapters. Blue arrows leading from one brain area to another index predictive processing streams, and red arrows represent prediction error signals. Blue arrows within brain areas stand for the facilitation/preactivation of predicted representations, and red arrows for later inhibition (in VC only).

**MECHANISMS OF PREDICTIVE PROCESSING**

A. prediction/preactivation  
B. prediction error

![Figure 30 - A cartoon model of predictive language processing in the brain](image)

At the presentation of a picture of a banana, predictive processing for the corresponding word “banana” is initiated (indicated by the thought bubble; 30A). This presumably triggers preactivation of the lexical-semantic representation of \{banana\}, housed in left temporal cortex, but sparing other lexical-semantic representations from lateral inhibition (see Figures 21-25). A predictive signal from the temporal lobe is
possibly sent to vmPFC, which in turn communicates the predicted representation to
VC (although direct connectivity between the temporal lobe and VC may also exist). The connectivity between each of these regions, i.e., the directionality of the arrows, remains to be investigated. In VC, predictive processing triggers the preactivation of the form representations associated with the lexical-semantic predicted representation of “banana,” causing lateral inhibition of irrelevant form features (see Figures 21-25). When the word “apple” is presented instead of the predicted “banana” (30B) error signals are communicated between brain regions in reverse order: First, in VC, inhibited form features need to be engaged, potentially accompanied by a simultaneous suppression of the predicted form representations. This process triggers an error signal to vmPFC, probably in addition to following the bottom-up processing stream from visual word form to lexical-semantic representation. In the left temporal lobe then, bottom-up activation processes lead to the activation of the lexical-semantic representation of “apple.” In the time between the lexical-semantic preactivation of “banana” and the lexical-semantic processing of “apple,” the activation level of neurons representing the predicted lexical-semantic representation of “banana” should have subsided a bit, enabling the selection of the representation of “apple” over “banana.”
Future directions

Clearly, some of the hypotheses derived from the adapted version of the Sensory Hypothesis will have to be tested in future research.

For example, I have not demonstrated the existence of predictive processing for syntactic representations. Moreover, it is likely that predictions may be generated for any level of analysis, including e.g., phonological, prosodic, or pragmatic information. I have also not been able to pinpoint the exact roles of ventro-medial prefrontal and left middle temporal cortex in predictive processing. In addition, while the minimum norm estimates presented in Chapter 8 are suggestive of a relationship between processes in ventro-medial prefrontal cortex and visual cortex respectively, the analysis used here does not allow us to demonstrate such a relationship directly: It can only be inferred from the temporal succession of effects. Further, the exact nature of ‘form features’ remains a mystery: Letter representations? Global word shapes? Multidimensional phonetic feature clouds? Future research should be aimed at elucidating the nature of these form features, for example by way of disrupting potentially relevant levels of form representation (e.g., case alternation).

And then there is the fact that participants were asked to match words to pictures, a task which does not exactly approximate the dynamic and complex nature of everyday language comprehension. In fact, one could question whether the experimental design that was used in Experiment 4 satisfies even basic concerns of ecological validity. This issue does not only apply to the experiments reported here:
Compelling evidence pertaining to the nature of predictive processing under natural circumstances is few and far between. Even the most convincing papers seem to have flaws: For example, in the study by Summerfield and Koechlin (2008) cited above, context (predictive vs. non-predicted) was confounded by task (a match/mismatch task in predictive contexts and a forced choice task in the non-predictive contexts). As a result, despite the growing number of scholars in the language processing research community that adopts predictive approaches, direct evidence that the brain engages in predictive processing in its natural habitat remains limited at best, and a careful evaluation is warranted of any finding that is termed as a predictive effect. I hope I have invited you to do exactly that.
BIBLIOGRAPHY


