Words and Rules in the Brain

by

Jaemin Rhee

B.A., Biochemistry
Harvard University, 1993

Submitted to the Department of Brain & Cognitive Sciences in Partial Fulfillment
of the Requirements for the Degree of

Doctor of Philosophy in Cognitive Neuroscience

at the

Massachusetts Institute of Technology

June 2001

Signature of Author:____
Department of Brain and Cognitive Sciences
May 1, 2001

Certified by:____
Steven Pinker
Peter de Florez Professor of Psychology
Thesis Supervisor

Accepted by:____
Earl Miller
"Class of 1956" Associate Professor
Chairman, Department Graduate Committee

[redacted]

Words and Rules in the Brain

by

Jaemin Rhee

Submitted to the Department of Brain & Cognitive Sciences
on May 1, 2001 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in
Cognitive Neuroscience

ABSTRACT

The Words-and-Rules theory (WR) posits that different mental processes underlie regular and irregular past tense formation: regular forms are rule-generated ('add *-ed*'), whereas irregular forms are retrieved from memory. These mental processes are hypothesized to engage distinct neural mechanisms. The goal of the present thesis was to localize and differentiate the neural substrates of regular and irregular past tense generation. Two neuroimaging techniques, magnetoencephalography (MEG) and functional magnetic resonance imaging (fMRI) were used to test healthy, right-handed subjects who were native speakers of English in a past tense production paradigm, in addition to a lexical access study. The results indicate that there is a dissociation in both the time course of activation and brain areas involved for the regular vs. the irregular past tense formation.

Thesis Supervisor: Steven Pinker
Title: Peter de Florez Professor of Psychology

Acknowledgments

I would like to acknowledge the contributions of Steve Pinker, my advisor, and Anthony Wagner, one of my committee members, who (more than) generously gave me his time, both at his office and at the scanner. And I would like to thank the rest of my committee members, Ted Gibson, Nancy Kanwisher, and Alec Marantz, for their support and tough questions.

I would also like to thank all of the members of WagnerLab, Lila Davachi, Matt Hutson, Dav Clark, Itamar Kahn and David Badre, without whose enthusiastic support I would still be languishing in pre-dissertation hell; Andrew Stringfellow, whose sense of humor helped me cope with the vagaries of MEG research; Morris Halle for introducing me to phonology and other such things; Boris Magasanik for just being a great friend and former tutor; William Marlsen-Wilson for moral support and encouragement from across the Atlantic; and (in no particular order) David Poeppel, Michael Ullman, Daniel Casasanto, Daniel O'Connor, TC Ulman, Duane Watson, Yaoda Xu, Albert Lee, Tony Harris, Kevin O'Connor, Jennifer Ganger, Ian Dobbins, Seppo Ahlfors, Yuri Ostrovsky, Liina Pylkkänen, Liu Jia, Fei Xu, Rutledge Ellis-Behnke, Dan Grodner, Tessa Warren, Denise Heintze, Gregg Solomon, the kids at No 6 club, Dr. HJ Friedman, and Marie Lamb for support, advice and very often just a good laugh.

Finally, I would like to thank all the people with whom I have been playing and making music, in particular my most wonderful chamber music coach, John Harbison; my piano teacher, Tatyana Yampolski; the very many Fluff Artists, and Yukiko Ueno, who have all helped to keep me sane during my years as a graduate student.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Paul Broca's seminal paper on aphasia, "Remarques sur le siège de la faculté de la parole articulée; suivies d'une observation d'aphémie (perte de parole)" appeared in 1861, in the midst of the phrenology debate (Caplan, 1986). The question Broca was trying to address was whether various brain functions could indeed be localized to specific brain areas, as the phrenologists argued. Based on the behavioral patterns of his patient Labourgne, who could comprehend, but not produce language, and a subsequent brain autopsy of the unfortunate man, Broca concluded that language production, but not comprehension, was localized in the left frontal cortex (Broca, 1861 and 1865). The brain area he pinpointed as necessary for language production, Brodman's area (BA) 44/45, has since been called "Broca's area" (see Fig. 1), and the disease accompanying a lesion in that area "Broca's aphasia".

Put Fig. 1 about here

Soon after, two other brain areas were implicated in language function, Wenicke's area (BA 22) and the Arcuate fasciculus (see Fig. 2). The standard model of language production and comprehension that came out of these earlier aphasiological studies is commonly referred to as the Wernicke-Geschwind model of language and gesture (see Fig.2).

Put Fig. 2 about here

The model proposes the following steps for the simple task of repeating a word: the information flows from the primary auditory cortex (BA 41) to higher areas in the auditory cortex (BA 42), and then to the angular gyrus (BA 39) which associates incoming sensory information. From there, the information passes on to Wernicke's area (BA 22) and then, on to Broca's area (BA 44/45) via the arcuate fasciculus. The perceptual processing of the perceived word occurs in Wernicke's area, and the grammatical and phonological encoding necessary to repeat/articulate a word in Broca's area. The latter information is then sent to the motor cortex, for articulation.

According to the Wernicke-Geschwind model, a lesion in Broca's area results in Broca's aphasia, where patients are severely impaired in language production, especially in the grammatical aspects, but retain language comprehension. A lesion in Wernicke's area results in the opposite clinical observation, Wernicke's aphasia, in which patients have fluent language output, but are impaired in language comprehension. And a lesion in the arcuate fasciculus would result in conduction aphasia, in which patients suffer from both object naming and word repetition deficits.

With the onset of modern cognitive and imaging studies, however, this standard model has been challenged on two separate fronts. First, simply mapping language functions to a few distinctive cortical areas has proved to be elusive. With respect to neuroanatomy, we now know that many other brain areas besides the original three are crucial, including subcortical areas, such as the basal ganglia and the brainstem (see e.g. Brunner et al., 1982, and Ullman et al., 1997b). Other brain areas that have been implicated in verbal processing are the anterior LIPC (Left Inferior Prefrontal Cortex, BA 45/47) for semantic selection (see e.g. Wagner, 2001) or BA 40, in the posterior parietal cortex, for short term storage and retrieval of verbal information (see e.g. Jonides et al., 1998; see chapter 6 for further discussion).

Second, the standard Wernicke-Geschwind model fails to explain all of the available data on the human language faculty. For instance, not all strokes that affect BA 44/45 result in Broca's aphasia, while some strokes that leave BA 44/45 intact do (see e.g. Basso et al., 1985) .

On the cognitive side, the number of “stations” involved in language processing has proliferated: at the very least, we now assume the existence of both phonological (auditory) and visual (graphemic) input and output stores, a lexical/semantic store, a grammatical processing station, a facility to interpret prosody, as well as working memory components (which have yet to be worked out fully).

Furthermore, with the onset of modern linguistic theory, language itself has been split into two different components: one is comprised of the rules that operate within a language, such as 'a reflexive pronoun like *himself* needs a referent within the sentence it occurs in'. The other is comprised of all the parts these rules work on, such as words (e.g. *want*), idioms (e.g. *a change of heart*) and morphemes (e.g. *-s*, as in *brother-s)*, which I will hereafter broadly refer to as 'words' from now on. This entails a separation between language processes which are based on rules, e.g. syntactic processing, and language processes which are based on words, e.g. lexical processing.

The work reported in this thesis tries answer the following questions: what are the neural computations regular and irregular morphology depend on? The particular case investigated was the English past tense.

Why the English past tense? Why not any other language, or any other construction? There are two reasons the English past tense is an ideal system to study language in general. The first is that the English past tense has been intensely studied in several disciplines, and offers a wealth of data and theoretical frameworks.

Second, it meets the requirement of being both simple enough to be tackled, and yet complex enough to be relevant. It is a rather straightforward system, with only two classes of verbs, regular and irregular. The regular pairs (such as *walk-walked*) are entirely predictable (i.e. add *-ed* to the stem), while the irregular forms (such as *bring-brought*) are not and have to be learned and committed to memory. Otherwise, they are exactly the same, as they are matched in meaning (pastness), grammar (tensed), and complexity (one phonological word). However, past tense formation also has grammatical components, much like the formation of a sentence, while being much simpler than sentence processing itself: the feature of 'past-ness' has to be looked up and incorporated, and a suffix has to be added to the end of the stem form, following a rule (past tense = stem + *-ed*). Similarly, during the formation of a sentence, tense features have to observed, and a strict order of the components has to be followed (e.g. the verb always follows the subject; see table 1).

Given the differences and similarities between regular and irregular English verbs, the question my thesis tried to look into should perhaps be sharpened and rephrased in the following manner: are regular and irregular English verbs categorically distinct, i.e. do they depend on only one, or on two distinct neural computations?

Obviously, two general theoretical positions are possible, one that states that regular and irregular verbs are not categorically distinct, and depend only one neural computation (Single System Theories). The other would state that regular and irregular verbs are categorically distinct, and would depend on two distinct neural computations (Dual System Theories).

The Single System Theory I am going to discuss briefly is called Connectionism. Connectionism takes much of its philosophy from the branch of AI called Parallel Distributed Processing, or PDP (Rumelhart & McClelland, 1986). In PDP, the parts of the brain relevant to a particular cognitive function is represented by a gigantic neural network, which goes about its business of cognition by two basic processes, summation of input from various input units, and setting of synaptic weights that determine how strongly the activation of one unit affects the next unit (see Fig. 3).

Put Fig. 3 about here

The above network has only one type of function, deriving the output from the inputs in some sort of additive fashion, and this function is distributed over the entire network (hence the name PDP). By its very nature, the network does not distinguish between different types of procedures, nor could it reserve particular areas for particular calculations.

Connectionist models (e.g. Rumelhart & McClelland, 1986; Rumelhart & McClelland, 1987; Hahn & Chater, 1998; Marchman, Plunkett & Goodman, 1997; Plunkett & Nakisa, 1997) propose that one generalized network subserves both regular and irregular forms. The difference between them is not due to a categorical distinction, but to the fact that they inhabit different probability spaces within the same network: there are many more regular verbs than irregular ones, which would induce the learner to assume that in most cases the past tense form is the stem + *-ed;* only in a few special cases would the past tense form look different. In other words, the pathways that lead from the input nodes to the correct output nodes are much more often traveled for regular verbs than for irregulars.

For irregular verbs, the network knows to derive *kept* from *keep* and *flung* from *fling* because they are similar to other verbs which undergo the same transformation: *keep-weep-creep-sleep* all become *kept-wept-crept-slept*, for instance, and *fling-sting-string* all become *flung-stung-strung*. Words such as *bring* and *eat* are problematic, since the information about their phonological resemblance is not useful. In the case of *bring*, the network might be tempted to derive the wrong forms *brung* (by analogy to *slung*) or *brang* (by analogy to *sing*). In the case of *eat*, the network might want to output *et* (by analogy to *meet-met* or *eat* (by analogy to *beat-beat*), instead of the correct *ate*. For words like these, the network has to be extensively trained before it can produce the correct forms.

The processing differences between regular and irregular verbs, then, are solely due to frequency, similarity effects and their distribution in phonological space, and not due to a categorical distinction between them (see Pinker & Prince, 1994; Seidenberg & Hoeffner, 1998; and Bates, 1999, and the references cited there for a more detailed discussion).

Assuming a *single* network which is not comprised of sets of smaller networks, this entails that there should be no *macroscopic* double dissociations between the computations underlying regular and irregular past tense formation. Microscopically, there should be differences in activation patterns for individual items, but these should not translate into large scale dissociations (see Bullinaria and Chater (1995 and 1996) for detailed arguments as to why connectionist models in general cannot display double dissociations (see also Pinker, 2001)).

The general problem with the connectionist models is that they have yet to implement a connectionist network which can successfully account for all the data on the English past tense production. Their models either fail to behave like English speakers, or if they do, can do so only with artificial manipulations (e.g. by altering the ratio of regular to irregular verbs the network encountered, see Pinker and Prince, 1988, and Marcus et al., 1992, for further details.)

Representing Dual System Theories is the Words and Rules theory (Pinker, 1991). In this framework, that regular and irregular past tense forms are different, in that they are stored, produced and processed differently (e.g. Pinker, 1991a; Pinker & Prince, 1988; Pinker & Prince, 1994). Regular past tense forms are usually derived by the application of a default rule, which is 'add *-ed* to the stem', whereas irregular past tense forms are stored in an associative type of memory and have to be recalled. Regular and irregular past tense formation should therefore be distinct in a fashion similar to the distinction between syntactic and lexical processing, and depend on distinct neural computations.

Note that the issue here is finding dissociating neural computations, and not just any differences as such. Clearly, regular and irregular verbs are different, in token and type frequency, in phonological neighborhood densities, in productivity, in difficulty. What is important is whether it can be shown if they rely on distinct neural processes or not, once confounds such as frequency and difficulty are removed, and previous research indicates that this is indeed possible.

Bolstering the Words/Rules theory, many other studies have concluded that a double dissociation can be observed between regular and irregular processing, even though the results are not always unambiguous. Among clinical studies, Marslen-Wilson & Tyler (1997) reported a double dissociation between two different clinical population using an auditory priming task. Priming in general refers to the facilitation of a behavioral response based on prior exposure to the material. A typical example is semantic priming, in which subjects are presented with a prime word, and then are asked to judge whether the following word, the target, is a lexical item or not. If the target word is semantically related, the reaction times are significantly shorter.

In Marslen-Wilson and Tyler's study, both prime and target words were presented auditorily, and subjects were asked to judge whether the target word was a lexical item or not. Their results are as follows: one aphasic subpopulation, which no longer showed semantic priming, also failed to show priming for irregular forms, while priming for regular forms was intact. In another subpopulation, the reverse was true: priming for regular forms was severely impaired, while both semantic and irregular priming were still intact.

Another double dissociation was reported by Ullman and his colleagues (Ullman et al., 1997a): Alzheimer's patients, who undergo severe memory loss, performed worse than control subjects when forming irregular forms. Parkinson's patients, on the other hand, who have difficulties with initiating movements, which Ullman et al. posit to be connected to the application of rules in general, showed the opposite pattern.

In addition, child acquisition data indicate that there is a difference between regular and irregular verbs. Around age two, children occasionally overregularize irregular verbs (*holded, eated*). This occurs when they start marking regular verbs for past tense reliably in obligatory contexts. As they get older, the frequency of their errors declines. Regular verbs are irregularized (*wipe/wope*) occasionally, but much less often (e.g. Pinker, 1991b; Marcus, Pinker, Ullman, Hollander, Rosen & Xu, 1992). This behavior can be mimicked by a connectionist network only if the ratio of irregular to regular verbs in the input training set is drastically altered, which does not correctly reflect the input children receive..

As for neuroimaging studies, the first neuroimaging paper published on the English past tense was Jaeger et al.'s PET study (Jaeger et al., 1996). They asked subjects to perform two tests. The first one was to view regular (e.g. *walk*), irregular (e.g. *teach*) and novel (e.g. *sitch*) verb stems on a screen, and the second to silently generate the past tense forms. Their results indicated that there was a distinction between the brain areas that subserve irregular and regular past tense morphology: area 46 (Left dorsolateral prefrontal area) was more active in the regular past production, than in the irregular past production. Area 10 (Left superior frontal gyrus), on the other hand, was more active in the irregular than in the regular past production.

However, there were several problems with their study, which called these clear-cut results into question. First, they did not perform the relevant subtractions. They only subtracted each reading task from its accompanying past tense task, but not regular from irregular past formation. The last subtraction is crucial, as they cannot compare regular and irregular past tense formation directly otherwise. Ideally, these kind of data should be analyzed as a subtraction of subtractions ((regular past minus regular read) minus (irregular past minus irregular read)). In ANOVA terms, that is equivalent to looking for the interaction term. Without the last comparison (the interaction term), the authors cannot claim that they found a dissociation between the two conditions, as the differences could just reflect type II errors..

Second, since the above was a PET study, they had to present their stimuli in a blocked fashion, i.e. they had to show blocks of verbs within the same category. Blocking stimuli always leaves open the possibility that subjects come up with a strategy to make the task easier. One possibility is that the subjects spent less attention on novel words and regular verbs less than on irregular verbs, since they knew with the first two categories that all they had to do was add *-ed* to the stem. In contrast, they had to pay attention to every irregular item since an irregular past tense form is not necessarily predictable (compare *sing-sang-sung* with *bring-brought-brought*).

Third, the brain areas Jaeger and her colleagues reported for regular (area 46) and irregular (area 10) past formation do not agree with the results of three other neuroimaging studies (Indefrey et al., 1997; Ullman, Bergida & O'Craven, 1997; Bergida, O'Craven, Savoy, & Ullman, 1997).

Indefrey et al.'s study was a crossed-design PET experiment of German verbal inflection; subjects were asked to overtly insert the inflected past tense or past participle form of a stem form into a sentence frame; the baseline condition consisted of reading the inflected, and not the stem forms. They reported that the right inferior temporal gyrus (BA 37) and the left angular gyrus (BA39) were more active for regular and for irregular verbs, and that, among others, left BA 46 and let BA 44/6 were more active for irregular verbs than regular verbs, in contrast to Jaeger et al.'s study.

Ullman et al. (1997) reported a blocked fMRI study, in which subjects were asked to silently produce the past tense forms of stem forms. The baseline condition was fixation, and not reading the stem. Their results indicated that left frontal cortex and the basal ganglia were more active for irregular than regular verbs. Furthermore, in temporo/temporo-parietal regions, irregular verbs showed a decrease in activation relative to fixation, whereas regular verbs did not show this decrease. However, in a later abstract (Bergida et al., 1998), they reanalyzed the same experiment and reported that temporo/temporo-parietal regions showed a decreased activation for irregulars but not regulars in, and "a left prefrontal region" showed an increase in activation for irregulars, but a decrease in activation for regulars.

The results of the three studies outlined above show little convergence. However, these studies all differ from each other in their experimental design and data analysis, so perhaps a direct comparison is somewhat inappropriate. As Seidenberg and Hoeffner (1998) concede that a true neural double dissociation would be strong evidence against a connectionist model; however, they also point out that the data reported so far are not necessarily convincing, and that serious confounds call the arguments in favor of Dual Systems in question. The experiments presented here are hopefully a first step toward showing clearly what the neural computations are that regular and irregular past tense formation depend on, and whether they show a clear double dissociation or not.

The starting point is a set of preliminary studies I conducted on the production of English past tense forms, using magnetoencephalography, or MEG (Rhee, Pinker & Ullman, 1999), which were subsequently fortified with functional magnetic resonance imaging (fMRI) experiments. The next 3 chapters will present a detailed past tense model (chapter 2), a description of the imaging techniques involved (chapter 3) and the data from 2 imaging studies of past tense formation (chapters 4 and 5).

Chapter 2

A Specific Past Tense Model

As stated in the previous chapter, the Words-and-Rules theory predicts a double dissociation between the neural computations regular and irregular past tense formation depend on, while connectionist models do not, at least not on a macroscopic scale. In this chapter, I will outline a past tense model that incorporates the Words-and-Rules theory, which can then be tested against the connectionist model.

The Words-and-Rules theory proposes the following differences between regular and irregular past tense formation: regular past tense forms are usually not stored, but are derived by adding the suffix *–ed* to the stem, following a 'default rule'. Irregular past tense forms, on the other hand, are stored in memory, with some associative properties, and have to be looked up (e.g. *teach-taught*). While the search for the correct form is going on, however, the default rule is activated as well. Successful look-up will suppress the application of the rule, and the output is the correct irregular form. If look-up fails, on the other hand, because the memory trace was not strong enough, for instance, or because there actually is no stored item in memory, then the rule is applied, as an option of last resort, or default. The output then consists of forms like *teached* instead of *taught* (memory failure for known irregular item) or *wugged* (lack of memory trace for a novel word). Note that this is a parallel model, in which lookup and activation of the rule occur simultaneously (see Fig. 4).

Put Fig. 4 about here

This distinction between regular and irregular processing hinges on the use of a 'default rule'. 'Default rule' here is a technical term. In essence, it means that irregular morphology is the marked form, which applies only in very narrowly prescribed circumstances, whereas the default is the unmarked form, and applies everywhere else. Among other things, the default past tense rule applies when the word is novel, i.e. has no entry in the lexicon and could not possibly be marked and so undergoes the default rule (*wug-wugged*).

Therefore, the relevant distinction between regular and irregular past tense formation in English would hold for other morphological processes that involve regular vs. irregular words, in any other language. To give another example: the plural formation of English nouns is mostly regular ('as the default rule, add *-s* to the stem'). Many common nouns, however, have irregular plural forms, e.g. *child-children, foot-feet*. These forms have to be stored in the lexicon and recalled, just like irregular past tense forms.

According to the theory, it is coincidental that in English, the class of verbs (and nouns) to which the default rule applies is much larger than the class of verbs (and nouns) to which the marked rules applies. In German, the class of verbs which undergo the default rule, which is 'add *–t* to the stem to form the past participle' as in *kaufen-gekauft*, is smaller than the class which undergoes the marked rule, which is 'add *–en* to the stem to form the past participle', as in *singen-gesungen* (Marcus et al., 1995).

Thus, the specific past tense model that is outlined below should also apply to other kinds of morphological processing, with some minor modifications. For purposes of clarity, however, I will couch the discussion in terms of the English past tense formation.

In overview, the basic steps of deriving the past tense forms are the same for both kinds of verbs: the word has to be heard/seen and then it has to be recognized/looked up in a sensory input lexicon (either visual or auditory), for perceptual recognition (Caramazza, 1997). This first step typically applies only to psycholinguistic experiments in which subjects are presented with the stem form and are asked to produce the past tense form. During voluntary speech, the stem would be accessed directly in the lexicon.

The next step is to determine the correct semantic, syntactic and morphological features of the verb, including its 'past-ness'. These have to be looked up in the lexicon. Regular verbs undergoing the default rule for past tense formation do not have to be marked, by definition, while irregular words are. 'Marked' here means that there is a connection between the stem form and the inflected form, which indicates that the desired form is to be found in the lexicon.

Once the verb has been properly processed in the lexicon, and its features have been correctly assigned, activation spreads to the phonological or the graphemic output lexicon, where the phonological or graphemic features of the verb are looked up. After this point, the pathways for regular and irregular verbs diverge. In the case of regular verbs, processing their past tense forms is straightforward. The default rule automatically kicks in and concatenates the stem with the suffix *-ed*, to produce forms like *walked*.

However, for irregular verbs, finding the correct past tense form is more complicated. The retrieval call to the past tense form of an irregular verb will activate both the default rule and the search for the stored past tense form. As the search for the correct irregular form proceeds and the activation of the memorized past tense form passes a certain threshold (i.e. it becomes clear that it can be found), the default rule has to be prevented from applying, otherwise the output could consist of incorrect forms such as *bleeded* (overregularized) or *bleded* (doubly marked). If the search for the irregular form is not successful, due to memory failure, for instance, the default will not be suppressed, and a form like *bleeded* is produced. Alternatively, an incorrect form such as *brang* could be produced as well as a consequence of memory failure. In this case, the regular rule was successfully suppressed, but the correct stored form was not retrieved, and *brang* was output by analogy to *sing-sang*.

Finally, after the verb has been successfully processed, either through the rule or by lexical lookup, activation spreads to the output (motor) areas, in preparation for writing or pronouncing the past tense form.

What is unclear is the role of verbal working memory. For both categories of verbs, verbal working memory could possibly be involved in keeping the stem form on line while the past tense form is produced. This seems unlikely, since most language functions seems to be highly automated, given the fast reaction times (mean RT=808 ms for irregular past tense formation and mean RT=780 ms for regular past tense formation in the experiment reported here, see chapter 4)[1].

In addition, there might be a selection component involved in the formation of the irregular past tense forms in order to choose the correct irregular form among different possibilities. This might be implausible, given how quickly irregular past tense forms are produced (mean RT=808 ms in the experiment reported here, see chapter 4). On the other hand, stem completion tasks, which obviously have a selection from memory component, can have mean RTs of 1045 ms (Buckner et al., 1995), which is not that much slower.

Adding specific brain areas to the general outline is difficult, mostly because our knowledge of localization of function in the brain is still preliminary. Below, I will present a model that is based on various research data coming out of studies on lexical access (e.g. Levelt, Roelofs & Mayer, 1998), aphasia (e.g. Ullman, Corkin, Coppola, Hickok, Growdon, Koroshetz & Pinker, 1997), and morphological processing (e.g. Koenig, Itzel & Caramazza, 1996),

The model is specific to my experimental paradigm, in which subjects were asked to pronounce the past tense form of a visually presented stem form. Therefore, I will only mention the visual input lexicon (for the visually presented stimuli) and the phonological output lexicon (for the production of the past tense forms), and omit the graphemic output and phonological input lexicons.

---

[1] Reaction times alone cannot decide whether a task is automated or not, of course; to my knowledge, no data have been reported on dual task studies involving past tense formation.

For both kinds of verbs, the pathway to a past tense form begins with the visual input lexicon, which could be stored in the visual cortex (BA 17, 18, 19). This occurs in the dominant hemisphere for language, usually the left (see e.g. Petersen et al., 1989, and Petersen et al., 1990). Alternatively, the visual input lexicon could be stored in the left occipito-temporal region, in the middle portion of the left fusiform gyrus, which is also known as the Visual Word Form (VWF) area (see e.g. Cohen, Dehaene, Naccache , Lehericy, Dehaene-Lambertz, Henaff, & Michel, 2000).

Then, the words have to be looked up in the lexicon, which is taken to be distributed over the perisylvian region. It has not been settled yet whether the lexicon should be located bilaterally or left-lateralized. There is considerable evidence for bilateral distribution, regardless of the subject's handedness (Pulvermüller, 1998; Martin, Haxby, Lalonde, Wiggs & Ungerleider, 1995; Martin, Wiggs, Ungerleider & Haxby, 1996; Damasio & Tranel, 1993; Caramazza & Hillis, 1991). On the other hand, Damasio and his colleagues (Damasio, Grabowski, Tranel, Hichwa & Damasio, 1996) reported that lexical retrieval activated only left temporal areas. Since this dispute is not settled, I will leave this particular point open.

Then the activation spreads to the phonological output lexicon, which is mostly localized to the dominant hemisphere, but should show some representation in the other hemisphere as well (Koenig et al., 1992), and which is tentatively located in the left Superior Temporal Gyrus (STG), or BA 21/22 (see e.g. Howard et al., 1992; Price et al., 1996a; Price et al., 1996b; Petersen et al., 1989; and Fiez and Petersen, 1998, for an overview). BA 21/22 has been consistently implicated in the above studies in reading words, both silently and aloud, and might be involved in the transformation from graphemic to phonological representations. In addition, it does not seem to be activated by non-word auditory stimuli (Lauter et al., 1985). Hence, it might be a candidate for the phonological output lexicon.

After the phonological features have been looked up, the past tense processing takes place. For regular verbs, activation spreads to a grammatical processing station where suffixation (concatenation) takes place. This is tentatively taken to occur in the left inferior frontal gyrus, BA 44/45, based on studies which imply BA44 in syntactic tasks. For instance, Embick et al. (1999) implicated BA44/45 to be more active in a syntactic processing task (monitoring word order mistakes) than in a non-syntactic task (monitoring incorrectly spelled words in grammatically correct sentences). Also, Caplan et al. (1996) reported that three patients with lesions in BA 44/45 showed impaired syntactic processing capacities, while Stromswold et al. (1996) showed that BA44/45 was more active during the processing of syntactically more complex sentences.

Irregular past tense forms, on the other hand, are being looked up in the lexicon. While the irregular form is being matched and retrieved, the default concatenation process (in BA 44/45) has to be terminated. One suggestion, put forth by Jaeger et al. (1996), is that this is done through BA 10, which was more active during the IrregPast-IrregStem subtraction condition than in the RegPast-RegStem subtraction condition in their past tense production study[2]; another

---

[2] This result is not completely convincing, though, since they failed to report the correct double subtraction conditions, as discussed above.

suggestion is that the inhibition has to proceed through the basal ganglia, via the striatum (Ullman et al., 1997).

For both classes of verbs, once their past tense forms are derived or found, the activation would spread to possibly the left precentral gyrus of the insula (Dronkers, 1996), and the output areas in the motor cortex (BA 4), which would then result in the overt pronunciation/writing of the words.

The above model, with the tentatively assigned brain areas involved, contains fewer features than what the theory would demand exactly: steps such as the assignment of the grammatical feature of 'pastness' and the suppression of the default rule are not fully discussed. Nevertheless, it provides enough details to base imaging studies on it.

Using a past tense production paradigm, one could expect to find the following activations, according to the above model. There should be more activation in left BA 44 (for the application of the default rule) for regular verbs than for irregular verbs. Irregular verbs, on the other hand, should show more activation of temporal areas, where the lexicon is taken to be located. Although regular verbs have to be looked up in the lexicon as well, searching for the memorized past tense form in addition to looking up the stem form should represent a 'double dip', which would result in greater activation for irregular past tense forms than regular ones. Both should show equal activation in the visual input lexicons, and BA 4 (motor output area) if speech activity is involved in the task (see Fig.5).

Put Fig. 5 about here

For connectionist models, the following predictions might hold: assuming that the single English past tense network is part of the lexicon, both regular and irregular past tense formation should activate temporal lobes. While individual items should show different activation patterns from each other, and irregular patterns should be more similar to each other than to regular patterns, no macroscopic differences should be observed.

The relevant data will be shown in the chapters 4 and 5.

Chapter 3

Brain Imaging Methods

This chapter will give a brief overview of the brain imaging techniques used in this thesis, magnetoenecephalography (MEG) and functional magnetic resonance imaging (fMRI). These techniques are by now well established in the cognitive neuropsychology community, and do not need lengthy introductions and explanations. For more details, see e.g. Lewis and Orrison (1995), on MEG, and e.g. the Visiting Fellowship in fMRI Handbook (2000) for fMRI.

Magnetoencephalography (MEG)

We know from Maxwell's equations that electric currents generate magnetic fields and vice versa; the geometric shape of the magnetic field generated by the electric current can be predicted by the so-called "right-hand rule: point your right thumb along the direction of the current flow, and the magnetic field lines will point in the direction of your curled fingers, i.e. they will lie in concentric circles around the current (see Fig. 6).

Put Fig. 6 about here

Since much of neuronal activity is the result of electric activity, it follows that neuronal activity generates magnetic fields, some of which will be measurable extra-cranially. And measuring these extracranial magnetic fields is the basis of MEG. Due to various biophysical properties, the neuronal activity that mostly contributes to extracranial magnetic fields comes from the dentritic potentials of cortical pyramidal cells. In addition, the head is a spherical conductor, which absorbs magnetic fields tangential to the skull. And since only pyramidal cells which are oriented in parallel to the surface of the skull produce magnetic fields with components perpendicular to the skull's surface, only those contribute to the signal measured by the MEG sensor (see Fig. 7). So all in all, the magnetic activity the MEG sensors register represents only a small fraction of the overall neuronal activity going on, the radial components of the magnetic field generated by pyramidal cells located in the sulci[3] (see Fig. 8 for a typical dipolar current pattern).

Put Fig. 7 about here

---

[3] The source of the signals EEG measure are not restricted to such a large degree.

Put Fig. 8 about here

How does the magnetic field generated by neuronal activity compare to other, familiar magnetic fields? It is exceedingly small, as one might have expected (see Table 2).

Put Table 2 about here

Measuring magnetic fields that are a few hundred femtotesla strong, if that much, is no mean task. For this purpose, the measurements are performed in a magnetically shielded room by extremely sensitive, superconducting sensors called SQUIDS (Superconducting Quantum Interference Devices). The magnetically shielded room is designed to exclude external, stable magnetic fields, such as the earth's magnetic field, as well as fluctuating magnetic fields. However, low frequency fluctuations, such as those created by nearby cables, are much less effectively excluded, and have to be subtracted from the data prior to analysis.

There are several advantages to using MEG over other imaging techniques, such as Event-Related Potentials (ERP) and Positron Emission Tomography (PET). In contrast to PET, MEG provides a temporal resolution that can realistically depict ongoing neuronal activity (milliseconds vs. seconds), and it is completely non-invasive. In contrast to EEG, there is no smearing of the signal due to the skull and tissue surrounding the brain, because the tissues distort electric but not magnetic fields. The result is that it is more feasible to localize the source of the neuronal activity using MEG data than using ERP data, even though it is by no means uncontroversial to do so.

Localizing the source of activity ('source localization') using MEG data is commonly stated as "the inverse problem", going from the distribution to the source[4]. An inverse problem is computationally ill-posed in the sense that there is no unique computational solution to it, as follows from Maxwell's equations. In order to constrain the search space of possible solutions, one usually makes three assumptions. The first is that the source of the activity is a discreet point source current dipole, i.e. that the source of the activity can be localized to a discrete

---

[4] Computing source localizations with ERP data faces similar problems. The algorithm for using EEG data is even more complicated than for MEG data, because of the smearing of the signal across the different tissues, and because ERP data are measured as potential differences between electrodes (i.e. additional data analysis has to be performed to turn the potential differences into absolute values before source localization can occur). Note that ERP papers, until recently, rarely report source localizations, in contrast to MEG papers, and that there is a trend to equate the position of the sensor with the brain area that sensor is sitting on top of.

point on the cortex. The other two are that the head is a sphere, and that this sphere is of uniform conductivity[5].

Armed with the above assumptions, a standard algorithm will take the following steps to calculate a source for the magnetic activity observed. First, a particular current dipole with a specific location, dipole orientation and strength is assumed. Then the algorithm computes the field this hypothetical current dipole, located within a sphere of uniform conductivity, would have generated at each sensor. Next, the algorithm calculates the mismatch term between the actual signal recorded at each sensor and the calculated, hypothetical value. This mismatch term is squared and summed over all sensors to generate an overall mismatch value. Then, the parameters of the hypothetical current dipole are changed, and the whole process is repeated iteratively until a best fit is found. The best fit here means a hypothetical dipole that results in the smallest overall mismatch term between itself and the observed data, and its location, orientation and strength parameters are taken to reflect the relevant neuronal currents.

How reliable the hypothesized current dipole is, depends on the assumptions that went into the algorithm to compute the dipole. The natural question to ask is how approximate are these assumptions?. The third, that the head displays uniform conductivity, is perhaps the least controversial one. Even though different kinds of tissue (skin, bone, cerebrospinal fluid, brain) have different conductivities, they can be modeled as concentric spheres which then can be summed up and modeled as one large sphere with uniform conductivity.

---

[5] That the head is a sphere is the most convenient assumption to make, since it makes the search for a solution to the forward problem the easiest.

Error estimates that compare simulations of non-uniform vs. uniform conductivities conclude that the error margins are about 1 cm (Cuffin, 1983).

Similarly, approximating the head as a sphere leads to inaccuracies of localization, especially in the temporal lobes, but the level of resolution still remains in the mm range (Cuffin, 1993).

The first assumption, however, that neuronal activity can be modeled as one current dipole, is problematic. Obviously, modeling cognitive activity with one point source (or even multiple point sources) at any given time will result in only very crude approximations of the regions involved. Ideally, then, one would want to model current distributions, rather than point sources.

One possible way of doing this has been under development in the laboratory of Anders Dale at the MGH NMR center in Charlestown. His program, Freesurfer, is designed to combine MEG and fMRI data into "fMEG". The structural MR data would provide the anatomical basis for localizing sources of neuronal activity, and the functional data from MRI would be used to constrain the search space for neuronal sources. They would replace the assumptions currently made in calculating current source dipoles. The end results of one's data analysis process using Freesurfer would resemble the traditional fMRI images which show areas of activation, and not single current dipoles, but with finer temporal resolution than fMRI, in the range of milliseconds instead of seconds (Dale et al., 2000). This approach, however, is controversial. For one, it smears activation over large areas of the brain. In addition, the method could misidentify the center of activation of a distributed source.

Another possibility would be using the software program BrainVoyager, which initially was geared toward MEG and EEG data analysis; recently, the makers, Brain Innovation B.V., announced a collaboration with SPM, one of the leading software programs to analyze PET and fMRI data.

functional Magnetic Resonance Imaging (fMRI)

fMRI takes advantage of two unrelated facts. The first one is that certain atomic nuclei have a magnetic spin that can be measured, after they have been aligned and locked in phase. The second is that blood contains both water and red blood cells, which carry oxygen.

The first fact is relevant for the basic physics behind fMRI, since the basis of fMRI is nuclear magnetic resonance (NMR), which essentially measures the amount of magnetic spin present in the sample (for a more detailed course on NMR, see Saunders and Hunt (1982)). Not all atomic nuclei display magnetic spin (only those with odd numbers of protons do) and for those that do, the amount of spin varies substantially

For brain imaging purposes, the relevant nucleus is the hydrogen atom, due to its abundance in biological tissue (as water), and the strength of its signal. And because different kinds of tissue contain different amounts of water, fMRI can easily distinguish between them based on the different amount of signal measured.

The second factor, that blood contains both red blood cells and water, is relevant for distinguishing between more and less active brain areas. This has to do with the well-known fact that active brain areas show increased blood flow (Roy and Sherrington, 1890). As blood flow increases, so does the local concentration of oxygenated blood; oxygen consumption, however, does not markedly rise, with the effect that there is more oxygenated venous blood circulating (which of course leaves the question open why there is increased blood flow in the first place - see Fig. 9).

Put Fig. 9 about here

It so turns out that de-oxygenated red blood cell has a magnetic spin, and that the oxygenated red blood cell does not (see Fig. 10).

Put Fig. 10 about here

The de-oxygenated red blood cells' local magnetic fields dephase the signal from the hydrogen atoms in the blood, among other bodily fluids, effectively decreasing the signal that can be measured. Consequently, venous blood in active areas, which carries more oxygenated red blood cells (and therefore fewer de-oxygenated cells) will give a stronger signal than venous blood in less active areas. This is the so-called 'BOLD (Blood Oxygen Level Dependent) effect' and forms the basis of modern fMRI techniques (see Fig. 11). Thus, fMRI is an indirect measure of neuronal activity: it gives an index of local blood flow and/or volume, which has a curious relationship to activity.[6]

Put Fig. 11 about here

Given the obscure relationships between local blood flow/volume, blood oxygenation levels and neuronal activity, interpreting an increase/decrease in the BOLD effect as an increase/decrease in neuronal activity requires a certain leap of faith. On top of it, there are three additional pitfalls. First, fMRI data are always subtraction data, i.e. one usually sees differences in activation between conditions and not the raw data themselves. The rationale behind this procedure is that all of the brain is active at any given point, and one could not discern which areas are significantly more active in one condition than in another just by looking at the activation for one condition alone, especially if the condition is a complex task. Rather, careful subtractions of conditions should reveal the brain areas of interest for a given cognitive task, assuming that cognitive functions can be decomposed into several individual steps. If task A comprises steps 1-5 and task B comprises steps 1-6 of a particular cognitive task, then subtracting task A

---

[6] Again, the relationship between blood flow and blood volume is not straightforward; for the remainder of the chapter, I will simply use flow/volume to stand for the relationship between blood flow and volume.

from B should yield only those brain areas that are involved in the sixth step of the cognitive task, assuming there are no interactions. Selecting the correct subtractions is crucial for isolating the relevant effects.

Second, the data are rarely viewed in their raw format. Instead, statistical maps are generated, using programs such as SPM99, which present only those voxels which survive a certain statistical threshold for a given subtraction, thereby introducing another level of interpretation between the data and the researcher. Also, the data are only as good as the statistical assumptions used in any given statistical package.

Third, collecting the brain activity following more than one stimulus is inherently difficult with fMRI, since the hemodynamic response takes many seconds, and not milliseconds. Thus, the hemodynamic response recorded for one stimulus overlaps with the response for the next. Separating out the different, overlapping components of hemodynamic signal is more difficult in blocked designs than in event-related designs, since in the latter, one can jitter the trial presentations and therefore decompose the signal more easily.

However, fMRI does offer advantages as well: it offers excellent spatial resolution without having to solve the inverse problem, and information about the activity of the entire brain, as opposed to cortical pyramidal cells which are oriented parallel to the skull surface (see above). Again, combining fMRI and MEG into "fMEG" should be a remedy for the poor time resolution offered by fMRI.

Chapter 4

Testing the hypotheses, part one: An MEG study

This was a straight-forward experiment to test the hypotheses: ask subjects to either read stems or produce the past tense forms of the same stems, while recording their brain responses continuously with a whole-head MEG scanner.

Materials and Methods

Participants

36 right-handed volunteers (26 male), aged between 18 and 35 years (with a mean age of 20 years), who were native speakers of American English, gave informed consent to participate in the experiment.

Materials

The stimuli used in this work were developed by Michael Ullman, and used in other studies of regular and irregular processing (Newman et al., 1998, Ullman et al., 1997, Bergida et al., 1997, Newman et al., 1999, Ullman, 2000). They were constructed with the goal of matching the groups of regular and irregular verbs as closely as possible on several dimension (syllable length, frequency, pronounceability).

Sixty-four pairs of regular (reg) and irregular (irreg) verbs were obtained, according to the following criteria. Each pair of irregular and regular verbs was matched four-way for frequencies of both stem and past forms, i.e. the frequencies of each of the four members in each pair (irregular stem, irregular past, regular stem and regular past) were the same, according to the Francis and Kucera (Francis and Kucera, 1982), Associated Press (AP, 1988) and Cobuild (Department of English, University of Birmingham, 1980) corpora. The frequencies ranged from 427/300 (number of occurrence) for the pair *take/ask* (most frequent) to 0 for the pair *breed/vie* (least frequent) according to the Francis and Kucera corpus.

A pair of words received the rating of 1 for goodness of match when the difference in the natural logarithm of their frequency counts was smaller than 0.5. If the difference was between 0.5 and 1, the word pair received a rating of 1.5. Word pairs with differences larger than 1 were rejected. Furthermore, the mean frequencies of irregular and regular verbs were compared by a paired t-test, and the p-values were larger than 0.1 ((mean(irregular)=3.05, mean(regular)=3.05, p=.24 for the FK corpus; mean(irregular)=6.02, mean(regular)=6.04, p=.83 for the AP corpus; and mean(irregular)=4.77, mean(regular)=4.67, p=.29 for the Cobuild corpus).

Similarly, each word pair was tested for similarity in the difficulty of pronounceability of the past tense form, based on similarity of the final cluster. Thirty-one of the 64 pairs were judged to be extremely well matched, i.e. they ended in the same configuration of consonants and vowels, such as /-pt/ for the pair *wept-whipped*. The other 33 pairs were only moderately well matched, i.e.

they did not end in exactly the same configuration of sounds, but were still judged to be equally pronounceable (e.g. *caught-played*).

Procedure

Participants were recorded for 2 sessions, each of which consisted of 128 trials. Each trial began with a fixation cross, which was presented for 700 ms. This was followed by the stimulus, which was presented for 300 ms, and then by a mask consisting of 6 asterisks. The mask was set to disappear when the subject started to speak, and the screen went blank for either 500, 700, 900 or 1100 ms (see Fig. 12).

Put Fig. 12 about here

Subjects were asked to read aloud the verb they saw on the screen (Stem Condition, session 1), or overtly produce the past tense form of the verb they saw (Past Condition, session 2; see Appendix I for the complete set of verbal instructions).

In addition, participants were asked to read aloud the 128 trial stimuli to become familiarized with them, and were given a brief practice session before recording. The verbs used for practice session were not used in the actual scanning.

During the experiment, the subject's head was positioned in a helmet-shaped Dewar and tightly pressed against its inner vault. The stimuli were presented on a Macintosh computer using PsyScope 1.2.5 (Cohen, MacWhinney, Flatt and Provost, 1993) and were projected via a collimating lens onto a mirror and from there onto the ceiling.

Within each session, regular and irregular verbs were randomly intermixed. Each subject was presented with all 128 verbs in each session. Across subjects, the stimuli were presented in a Latin square design (see Appendix II for a complete list of stimuli).

Data Acquisition

During the experiment, subjects lay in a dimly lit magnetically shielded room in the KIT/MIT MEG laboratory and neuromagnetic fields were measured using an axial gradiometer whole-head system (Kanazawa Institute of Technology, Japan). After the participation of the first 21 subjects, the KIT system was upgraded from 64 channels to 93 channels.

Head position with respect to the MEG sensor array was measured with 5 head position indicator coils placed on pre-defined scalp sites. In addition, a head frame coordinate system was established by recording the coordinates of the LPA, RPA, and the nasion of the subject, and the positions of the 5 head position coils with respect to these three landmarks were recorded as well. Later on, these two frameworks were combined in a so-called coregistered probe in the following manner: the coordinates of the subject's LPA, RPA and nasion were used to determine the point of origin and to set up a spherical 3D coordinate system. The position of the 5 head position coils in this framework were compared to their position in the MEG sensor framework, which then enabled the positioning of the MEG sensors within the coordinate system set up by the subject's landmarks.

At the beginning of the recording session, the magnetic signals produced by the head position indicator coils on the scalp were measured by the sensors to obtain head position with respect to the sensor array.

Data were recorded at 500 Hz, with acquisition between 1 and 200 Hz. The recording time for each session varied between 12 and 15 minutes, depending on how quickly the subject responded.

Behavioral Data

The reaction times were calculated from the onset of the visual stimulus. Incorrect responses, and RTs deviating over 2SD from the mean for the particular subject were rejected. The same trials were excluded from the MEG averages. The averages across subjects were then compared with repeated measures ANOVA with factors of task, subject and regularity.

Signal Analysis

The raw signals were first processed to remove contamination from the third rail of the subway[7]. Before the laboratory upgrade, noise-reduction consisted of subway noise subtraction using a signal-space projection method; after the upgrade, the Continuously Adjusted Least-Squares Method was adopted (CALM, Adachi et al., in press).

Then the signals for each condition were averaged separately off-line from 100 ms prior to onset of stimulus to 800 ms after the onset of the stimulus. Epochs containing MEG signals exceeding 2,500 fT/cm were omitted (<5% of all epochs), since they were assumed to include motion artifacts. In addition, the

---

[7] The third rail carries the current for the subway train to operate, and causes a very low frequency disturbance (around 0.5 Hz) of the recorded data.

averaged files were bandpass filtered between 2 and 40 Hz, and baseline corrected using the 100 ms pre-stimulus interval. The processed data were then displayed as isocontour maps (see Fig. 13 for a sample isocontour map).

Put Fig. 13 about here

Further data analysis consisted of RMS (Root Mean Square) analysis for source analysis (see below). The sources of the magnetic fields were modeled as equivalent current dipoles (ECDs) whose three-dimensional locations, orientations, and current strengths were estimated from the measured and averaged signals.

RMS analysis

After the data were cleaned of subway noise and averaged across trials within each subject, they underwent RMS analysis. In this kind of analysis, first a suitable subset of channels and a time window of analysis are chosen, then the data at each time point within the selected time window are squared for each channel and summed across all the relevant channels (see below for a description of how this is done). After this, the square root is taken, and the resulting data are plotted as a graph (see Fig. 14 for a sample RMS graph). This is done for each individual subject. The data, in the form of peak (or component) latencies and peak amplitudes, are then compared across subjects.

Put Fig. 14 about here

The RMS results give an indication of the overall activity across the selected channels, and can be used to compare one set of channels across different conditions. This method has the advantage of increasing the signal-to-noise ratio through the summation across channels. The disadvantage is that 'channel' can no longer be used as a factor in a statistical analysis, and that within-subject comparisons are not possible.

The selection of sensors is crucial. Typically, the selection is made by inspecting the isocontour maps for dipolar field patterns (see Fig. 7 in chapter 3) over the entire time course of the averaged data. If such dipolar field patterns occur, the channels selected are the ones that make up this particular field pattern (see Fig. 15 for an example of the selection of channels).

Put Fig. 15 about here

These channels are appropriate only for the time window which contains the dipolar field patterns based on which they were picked in the first place. If a later dipolar field pattern is distributed over different channels, the RMS is henceforth computed over the new set of channels. Usually, an averaged data set shows several dipolar field distributions during different time windows over its entire time course, often distributed over different sets of channels, and RMS is computed separately over each different set of channels.

Normally, only one dipolar field pattern is clearly discernible in any given time window, unless the concurrently occurring field patterns show bilateral and symmetric distributions. Otherwise, the field patterns would overlap, given that a typical field pattern takes up about 10-20 channels (64 channel system) and 15-25 channels (93 channel system) and there are only 32 (or 46) channels in each hemisphere. Overlapping dipolar field patterns result in field patterns which are no longer clearly recognizable as dipolar.

As stated above, RMS data have two indices, peak latencies and peak amplitudes. These peaks, in analogy to the ERP literature, are named after the typical post-stimulus onset time at which they occur, preceded by an M (as in 'MEG'). In contrast to ERP, all peaks are positive (by necessity, given how they were derived[8].) Based on the kind of cognitive task the subject is asked to perform, and the modality of stimulus presentation, one can expect to see certain peaks, which have been established by prior MEG studies. For a visually presented language task, the following peaks should be expected: M100, in the time window from 80 to 100 ms after onset of stimulus, M170, in the time window of 150-200 ms, M250 , in the time window of 230-280 ms, and M350, in the time window of 300-400 ms.

M100 and M170 represent components of early visual activation and occur bilaterally; localization of these components is tricky, for several reasons (see Ahlfors et al., 1999, for a discussion of localizing dipoles in the visual cortex.) The later two peaks, M250 and M350, are both taken to be related to lexical access and processing (Koyama, Kakigi, Hoshiyama and Kitamura, 1998; Kuriki, Takeuchi and Hirata, 1998; Kuriki, Hirata, Fujimaki and Kobayashi, 1996; Sekiguchi, Koyama and Kakigi, 2000; Helenius, Salmelin, Service, Connolly, 1998, 1999; Pylkkänen, Stringfellow, Flagg and Marantz, 2000; Embick, Hackl, Schaeffer, Kelepir and Marantz, to appear).

For the current experiment, then, only the later two established peaks, and peaks following them, were of interest, and RMS analysis was performed only on peaks occurring after 200 ms after onset of stimulus.

The number of channels used varied from 10 to 20 for the 64 channel system, and from 25 to 35 for the 93 channel system, depending on the time window and the dipolar field patterns they represented. The channels were all

---

[8] However, the magnetic field patterns still show a particular polarity and orientation.

located in the left hemisphere, as the activation patterns in most subjects were left-lateralized.

Subjects who did not show clear visual M100 and M170 peaks in their RMS data were excluded from the across-subject analysis; since the M100 and M170 represent early visual responses, their absence called the validity of the more downstream peaks into question.

Source Modeling

ECD analysis is typically performed if the RMS data and the sensor-by-sensor comparison show a significant difference between conditions over one specific set of channels , as confirmed by a repeated measures ANOVA design across subjects. The significant differences seen in these analyses go to confer a measure of validity on the results from the ECD analysis, in the following manner: the same channels, in the same time windows that were used for perform RMS analysis are used to do ECD analysis. The dipole localization algorithm will locate a dipole within the dipolar field pattern and within the time window observed; any difference in dipolar localizations can then be assumed to represent a statistically significant difference in brain activity.

The ECDs that best explained the most dominant signals were determined by a least-squares search, as outlined in chapter 3, using the EMSE dipole software package (EMSE Suite, Source Signal Imaging, Inc[9]). For each subset of channels, ECDs were calculated over the entire time window containing the dipolar field pattern. Only ECDs accounting for >80% of the field variance at selected periods of time for each subset of channels were accepted.

The output of the ECD analysis is a set of dipoles and their coordinates; these coordinates are located in a 3-D space determined by the subject's coregistered probe which was derived from the MEG sensor framework and the subject's head framework (see above, under 'Data Acquisition'). To plot the dipoles, one last step was necessary. The subject's coregistered probe had to be transformed from EMSE coordinates into Talairach space (Talairach and Tournoux, 1988) by matching the subject's landmarks with the standard ones. This transformation then was used to transpose the coordinates of the dipoles into Talairach space as well. Some distortions could not be avoided in the process, as a particular individual's landmarks do not necessarily match up well with the standardized ones.

Finally, the analysis was performed on both individual subject and grand averaged (averaged across sublets) data. For grand averaged data, one particular subject's coregistered probe was used, as no grand averaged probe was available.

---

[9] EMSE was previously calibrated by correctly localizing the auditory response to a 1kH tone test to approximately the auditory cortex.

Statistical Analysis

The latencies and amplitudes of the peaks of the RMS analysis were compared with ANOVA with repeated measurements with factors of task, regularity, subject and time, with subjects as the random factor, using StatView.

In addition, to further validate the ECD data for the past conditions, a sensor-by-sensor comparison was done, using the same channels as those for the RMS analysis, with a repeated measures ANOVA with factors of subject, regularity, sensor and time. The factor of 'time' did not represent individual time points, but time points averaged across a window that contained a peak, which is standard procedure in ERP research.

Predictions

For the Words-and-Rules model, the following predictions were made: both irregular and regular past tense formation should activate temporal areas (lexical lookup), but only regular formation should activate frontal areas (rule processing) as well. For the connectionist model, no neural dissociations were expected.

Results

Behavioral Data

As Table 2 shows, the reaction times for the Stem condition (mean=564 ms) were faster than for the Past Condition (mean=795 ms; F=173.6, p<.0001, df=1, 35). Within the Stem condition, the reaction times for irregular (mean=577 ms) and regular verbs (mean=560 ms) did not differ significantly, although they were somewhat faster for regular verbs (F=3, p=.09, df=1,35). Within the Past condition, the reaction times for irregular verbs (mean=808 ms) were significantly slower than for regular verbs (mean=781 ms; F=7.007, p<.05, df=1,35; see Table 3). The task X regularity interaction term was not significant (F=.419, p=.52, df=1,35)[10].

Put Table 3 about here

---

[10] This could be somewhat problematic, since there should be no differences between the two stem conditions, according to the Words-and-Rules theory. However, there are studies that indicate that the size of a derivational family can influence lexical retrieval times: the larger the family, the slower the retrieval (Baayen, 1997). Irregular verbs have a larger derivational family than regular verbs, since the psat tense forms are memorized. Also, there could be competitive inhibition from the memorized irregular past tense forms in the retrieval of the irregular stem forms.

Isocontour maps and RMS data

Figure 16 shows the relevant isocontour maps in each of the time windows for each of the conditions. These maps were chosen to best represent the dipolar field patterns occurring in each time widow for which RMS analysis was performed. All sensors are shown.

Put Fig. 16 about here

The following dipolar field patterns are shown: there is a visual response early on, around 100 ms (frame 1), followed by a another visual response around 170 ms (frame 2), followed by a field pattern around 230 ms (frame 3). The next one occurs around 320-370 ms (frame 4), and the last one occurs around 410-440 ms, and is confined to the past conditions.

Two sets of RMS analyses were conducted in the following time windows, based on the above dipole field distributions: the first lasted from 200 ms to 300 ms, and consisted of the channels distributed medially over the left hemisphere. The second lasted from 300 to 500 ms, and consisted of channels distributed more anterior to the first set (see Appendix II for a full list of channels).

The RMS data are as follows. For the first RMS set, from 200 to 300 ms, the first relevant peak occurs around 230 ms and is followed by one around 290 ms. For these two peaks, the following pattern of means can be observed: the latencies for the past conditions are significantly longer from those for the stem conditions (F=47.675, p<.001, df=1,30). Within the past and stem conditions, the latencies are not reliably different between regulars and irregulars (F=3.462, p<.1, df=1,30). The regularity X task interaction term is not significant (F=0.073, p<1, df=1,30). This is as expected, since the first two peaks occur too early to be involved in the past tense formation (see Table 4).

However, the amplitude data are unexpected. For both peaks, within the past condition, the amplitudes for the regulars were consistently smaller than for the irregulars. Within the stem condition, the amplitudes for the regulars were consistently larger than for the irregulars. As a result, the interaction term regularity X task is significant, at p=.001 (F=13.406, df=1,30; see Figure 17 for a graph for the amplitude data). This difference is puzzling, especially since the amplitudes for the later peaks show no significant interaction (see Table 5 for the RMS statistics).

Put Table 4 about here

Put Table 5 about here

Put Figure 17 about here

For the second set of RMS analysis, from 300 to 500 ms, the data are as follows: the next peak occurs around 320-370 ms, and both stem conditions show significantly faster latencies than the past conditions (F=54.003, p<.05, df=1,30). Within the past condition, the regular condition shows a significantly faster latency than the irregular condition (F=3.36, p<.001, df=1,30). Since this is the first time the irregular past and regular past conditions differ in their peak latencies, this could signify that the onset of past tense formation as such (as opposed to lexical lookup etc.) occurred sometime around 344 ms (which is the mean latency for the past condition), and that the peak following this one reflects the processing differences in regular vs. irregular past tense formation.

As stated above, there are no significant differences in the amplitudes (F=0.061, df=1,30, p<.5).

The final peak occurs around 410-440 ms, and is confined to the past condition. There is a significant difference in the latencies between the regular (mean=411 ms) and irregular conditions (mean=431 ms; F=26,96, df=1,30, p<.0001), but not in amplitudes (F=0.121, df=1,30, p<.5; see Tables 4 and 5).

In sum, the RMS data show that there is a significant and consistent difference between the stem and the past conditions in peak latencies, and between the regular past and the irregular past conditions, again in peak latencies. The RMS data were then used for the ECD data, which was expected to mirror the RMS differences within the various conditions.

ECD data

For single subject data, ECD analysis could not be performed successfully, presumably because the signal-to-noise ratio was too low, given the small number of stimuli (at most 64). Instead, ECD data from data averaged across subjects (grand-averaged) will be shown below. These data have to be taken with caution, for several reasons. First, dipoles derived from grand averaged data are much less interpretable than ECD data from single subjects, since averaging across subjects only means averaging across the data points at each channel, without taking differences in shape, general anatomy and relative head position with respect to sensors into consideration.

Second, I used a particular subject's coregistered probe (see above) to locate the dipoles, since I did not have an averaged coregistered probe at my disposal. Third, the transformation of the EMSE coordinates into Talairach coordinates resulted in some unavoidable distortions. Fourth, since the error margins for dipole localizations using a uniformly conducting sphere as model are under 1cm, only dipoles which are at least 2 cm apart in all coordinates could be safely deemed as distinct.

With all the above caveats, the ECD data are shown below, with the suggestion they could nevertheless give an idea of localization of activity (see Table 6 for exact coordinates and Figure 18 for an approximate localization.)

Put Table 6 about here

Put Fig. 18 about here

The distribution of dipoles (ECDs), from the above table, is as follows: there is no difference between the dipoles within the stem condition between regular and irregular verbs. Dipoles were found from about 200 ms on to about 340ms. Within the past condition, the dipoles for the regular condition started out together with the ones for the irregular condition, and were very similar from 200 ms to about 300 ms, again confined to left temporal areas. Thereafter, there was a split, as the two remaining regular dipoles between 300 ms and 500 ms moved anterior and occurred earlier, while the two remaining irregular dipoles remained in temporal areas, and occurred later than regular dipoles. This split can be taken to be reliable, since the distance between the regular and irregular sets is more than 2 cm, under the assumption that the error margin is 1 cm (Cuffin, 1985).

The localization of the dipoles to specific brain areas is left vague for the reasons outlined above; the Talairach coordinates for the late regular dipoles are approximately [-50, 6, 22], which could be taken to be located in frontal areas, even with an error margin of 1 cm, for the following reason: The frontal cortex is defined to be anterior of the central sulcus, the y coordinates of which range from about +35 mm to +12 mm. At z=+22 mm, the y coordinates of the central sulcus are about +15 mm. Even with an error margin of 10 mm, the dipole is still located forward for the central sulcus and therefore in the frontal cortex.

The posterior dipoles, whose Talairach coordinates are about [-40, -20, -10], are approximately located in temporal/parietal areas, given the error margin.

The temporal distribution of dipoles closely reflects the results from the RMS analysis. This is so by necessity, since the channels used for ECD analysis were selected on the basis of the RMS results. This procedure naturally biases the outcome of the ECD analysis; however, since the RMS results were found to be significant by objective criteria (repeated measures ANOVA), any difference found through in the ECD analysis can be taken to be valid, by the same criteria (see above.)

Affirming the differences in the regular and the irregular past dipoles, the sensor-by-sensor analysis revealed that the following interaction terms were significant: sensor X time (F=1.86, df=48,1440, p=.004), regularity X sensor (F=3.005, df-16, 480, p<.001) and sensor X regularity X time (F=2.385,, df=48,1440, p<.001; see Table 7).

Put Table 7 about here

Discussion

The data show the expected peaks that have previously been found in MEG studies of visual language tasks (M100, M170, M220, M350), in addition to another peak around 400 ms post-stimulus. M100 and M170 are connected to the visual processing of stimuli, and M220 and M350 possibly to lexical access; the last peak could then be related to the past tense processes.

The data also indicate that there is indeed a dissociation between regular and irregular past tense formation, both in time and space, in support of the Words-and-Rules theory over connectionist models. Irregular past tense formation takes significantly more time than regular past tense formation, as seen in the behavioral (RT) and MEG data. In space, regular past tense formation seems to involve anterior (perisylvian) regions, while irregular formation does not, according to the ECD analysis performed on grand averaged data. The anterior areas for the late regular dipoles could be the areas involved in the application of the rule (add *–ed* to the stem). The posterior areas for the irregular and early regular dipoles could be indicative of lexical processes (lexical look-up for both kinds of stems, lexical lookup of the irregular past tense form), since the lexicon is taken to be located in temporal areas (see chapter 2).

So far, this represents evidence to corroborate the Words-and-Rules theory, with the caveat that ECD analysis done on grand averaged data is less interpretable than ECD analysis done on single subject data.

Chapter 5

Testing the hypotheses, part two: An fMRI study

This was a replication of the past tense production study done with MEG. The goal was to provide better spatial resolution for the same past tense task than the ECD data analysis performed on MEG data.

Materials and Methods

Participants

21 right-handed volunteers (12 male), aged between 18 and 35 years (with a mean age of 20 years), who were native speakers of American English, gave informed consent to participate in the experiment.

Materials

The stimuli consisted of the same 64 pairs of regular and irregular verbs as in the MEG experiment.

Procedure

Participants were scanned for 1 session. The stimuli were presented for 700 ms, followed by 700 ms of a central fixation cross, followed by 100 ms of blank screen to indicate the end of each trial. Subjects were asked to either read silently the verb they saw on the screen (Stem Condition), or silently produce the past tense form of the verb (Past Condition; see Appendix III for a description of the verbal instructions). Each Stimulus was preceded by a cue word (either 'Read' or 'Past') which lasted for 300 ms, to indicate which task was required, and a fixation point, which lasted for 200 ms (see Fig. 19.)

Put Fig. 19 about here

Participants were given brief practice before scanning. The stimuli used during the practice were not part of the actual stimulus set used during scanning.

The stimuli were split into 2 subgroups of 32 pairs of verbs; each subject saw one group in the Stem condition, and the other group in the Past condition. The stimuli were presented in a pseudorandom fashion such that the four conditions (StemReg, StemIrreg, PastReg, PastIrreg) followed each other equally frequently to maximize the ability to decompose the hemodynamic signal associated with the different conditions. To add up to an overall scanning duration of 384 seconds for the scanning session, which was again calculated to maximize the BOLD signal contrast between the conditions, the stimuli, which added up only to 256 seconds (since there were 128 of them lasting two seconds each), were padded with a control condition. In this control condition, the visually presented stimulus was not a word, but a fixation cross, and the cue word was 'Fixate'. This condition also served as the baseline subtraction condition. In addition, the assignment of the subgroups was counterbalanced across participants.

FMRI methods

Scanning was performed on a 1.5T Siemens Sonata MRI system using a whole-head coil. Functional data were acquired using a gradient-echo echo-planar pulse sequence (TR=2 sec, TE=40 ms, 21 axial slices, 3.125 x 3.125 x 5 mm, .1 mm inter-slice gap, 192 volume acquisitions). High-resolution T1-weighted (MP-RAGE) anatomical images were collected for anatomical visualization. Head motion was restricted using a bite-bar apparatus. The stimuli were presented on a Macintosh computer and were projected via a collimating lens onto a screen, which was viewed through a mirror attached to the head coil.

Data analysis

Data were processed using the SPM99 (Statistical Parametric Mapping) software package (Wellcome Dept. of Cognitive Neurology, London). SPM99 goes roughly through three basic steps of data processing, preprocessing, model estimation and model inference. During preprocessing, images are first corrected for differences in slice acquisition timing by resampling all slices in time to match the first slice, followed by motion correction across all runs to correct for subjects' head movements. Subjects with head movements of more than one voxel size (3mm) were excluded from further analysis

Structural and functional data were then spatially normalized to the space of a standardized T2*-sensitive Echo Planar Image which is based on a representative image of an averaged brain of the MNI (Montreal Neurological Institute) stereotactic space, using a 12-parameter affine transformation along with a nonlinear transformation using cosine basis functions, to enable reliable localization of activation within subjects, and between subject comparisons.

Images were resampled into 3mm cubic voxels and then spatially smoothed with an 8-mm FW-HM (full-width, half-maximum) isotropic Gaussian kernel to increase the signal-to-noise ratio.

The next step is model estimation, during which the functional data are compared against an standard hemodynamic response, the activation one would expect to see if the null-hypothesis were correct . This is done using the general linear model in SPM99. Trials from each condition are modeled using a canonical hemodynamic response. Effects (the goodness of fit between the model and the actual data) are estimated, with session-specific effects and low-frequency signal components treated as error (for more details on SPM99, consult their webpage, http://www.mailbase.ac.uk/lists/spm).

The last step is drawing inferences from the data. This is done using linear contrasts, to obtain subject-specific estimates for each effect. These estimates are entered into a second-level analysis treating subjects as a random effect, using a one-sample t-test against a contrast value of zero at each voxel. Regions were considered reliable to the extent that they consisted of at least 5 contiguous voxels that exceeded an uncorrected threshold of $p < .001$. The maxima of these regions are localized on the normalized structural images and labeled using the nomenclature of Talairach and Tournoux (1988) and Brodmann (1909).

Finally, ROI (Region of Interest) analysis was performed, using the SPM ROI toolbox. ROI analysis was performed for two reasons, to extract the timecourse of activation for each of the condition, or to test the effects with increased power, due to pooling of multiple voxels across clusters.

In ROI analysis, regions (clusters) of interest are pre-selected, then the voxels in each cluster are collapsed. Finally, each individual subject's data set is re-sampled for those ROIs, their timecourse data extracted, and the extracted data can then be subjected to a repeated measures design ANOVA, with the random factor of subject, to isolate areas which show the desired significant interaction effects.

For this study, the goal was to explore for each ROI, which had a radius of 6 mm and at least 5 voxels, whether there was a significant interaction between regularity (irregular-regular) and task (past-stem), equivalent to a subtraction of subtractions. ROIs or voxels which showed a significant interaction could be taken to be more active either specifically in the irregular past formation condition ((IrregPast-IrregStem)>(RegPast-RegStem)) or vice versa.

One can choose ROIs either based on previous literature, or on one's own group data. In the latter case, one starts with the most general and unbiased contrast. If none of the ROIs show a significant interaction, one can go down to the less general contrasts. This is more likely to yield clusters with significant interaction effects, since the ROIs in these conditions were picked from data which had the bias built-in through a specific contrast analysis.

In this study, the most unbiased and general contrast condition was All-Fixation, which showed the areas active in all four conditions, from which the baseline fixation condition was subtracted. The next less general contrast condition was past-stem, in which both past conditions were collapsed and compared to both stem conditions, to look for areas more active in the pas conditions than in the stem conditions. The next less general contrast condition were the two single subtraction conditions, RegPast-RegStem and IrregPast-IrregStem. The RegPast-RegStem condition was analyzed to locate areas active in the regular past condition, irrespective of irregular verbs, and IrregPast-IrregStem, to locate areas active in the irregular past condition, irrespective of regular verbs. Finally, the most biased contrasts were the two double subtractions, (RegPast-RegStem)-(IrregPast-IrregStem) and vice versa (see Table 8).

Put Table 8 about here

Predictions

Based on the Words-and-Rules model, the following predictions were made: both irregular and regular past tense formation should involve temporal areas (lexical lookup), but only regular formation should involve frontal areas (rule processing). The connectionist model predicts that there should be no dissociations.

Results

All-fixation

The general pattern of activity for this condition, which indicates the general pattern of activity for the entire experiment, shows that most of the active clusters are limited to frontal and parietal areas. There is very little occipital and temporal activation[11].

This comparison did not yield any clusters or ROIs which showed a significant effect for the regularity X task X time interaction (see Fig. 20).

---

[11] This pattern holds for the conditions in which fixation was subtracted from only one experimental condition, e.g. IrregStem: activation was limited to parietal and frontal areas. These data are not presented, as they could be confounded with word effects, since the control subtraction condition was fixation, and not the same set of words processed in a different manner. In other words, the relevant effects could be due to the nature of the words themselves, and not due to the way they were processed.

Put Fig. 20 about here

Past-Stem

The active voxels in this contrast were mostly localized in the left frontal and parietal lobes, with distinctly less activation than in the all>fixation condition (see Fig. 21). This comparison did not yield any clusters or ROIs which showed a significant effect for the regularity X task X time interaction, either

----

Put Fig. 21 about here

Single Subtractions

Regular Past-Regular Stem

Since this contrast analysis was a rather biased one, one could have expected to find ROIs which show a significant interaction term. The regions exhibiting greater event-related responses to the Regular Past than to Regular Stem conditions were largely confined to the frontal lobes of both hemispheres (see Table 9 and Figure 22).

Put Table 9 about here

Put Fig. 22 about here

Given the hypotheses to be tested, the cluster I focused on was the one centered around the voxel with the coordinates [-48, 6, 18], which corresponds to BA 44/45, or Broca's area. Given that Broca's area has been implicated in a number of language related tasks, especially syntactic processing (see chapter two), this cluster might be involved in the application of the default past tense rule, i.e. 'add *–ed* to the stem'.

The other clusters are distributed in the following manner: one cluster is located in the right Superior Frontal Gyrus (SFG), or BA6, and one in the left SFG. One cluster is located in BA 7 (posterior parietal), and the other 4 clusters are located in the white matter.

However, none of the ROIs showed a significant effect for the regularity X task X time interaction.

Irregular Past-Irregular Stem

The regions exhibiting greater event-related responses to the Irregular Past than to Irregular Stem conditions were again largely confined to the frontal lobes of both hemispheres, and to one cluster in the left temporal lobe (see Table 10 and Fig. 23).

Put Table 10 about here

Put Fig. 23 about here

In addition to temporal lobe activation, the formation of irregular past tense forms involved the same area as the regular past tense formation, BA 44/45, on top of one other frontal area, centered around [-42, 33,-3], which corresponds to BA45/47 (see Fig. 22). A direct comparison between the clusters that are relevant to the two single subtraction conditions revealed why BA44/45 did not show a significant interaction effect: both conditions show active clusters in this area which are contiguous and overlapping (see Fig. 24).

Put Fig. 24 about here

Double Subtractions

Regular-Irregular

There was only one region exhibiting greater event-related responses to the Regular subtraction condition than to the Irregular subtraction condition, centered around [21, -33, 42], in the right parietal region (see Table 11 and Figure 25).

Put Table 11 about here

Put Fig. 25 about here

Irregular-Regular

There were 7 regions exhibiting greater event-related responses to the Irregular subtraction condition than to the Regular subtraction condition (see Table 12 and Figure 26.)

Put Table 12 about here

Put Fig. 26 about here

The significant clusters in this double subtraction condition can be described as follows: one cluster of 12 voxels, located in the anterior left inferior prefrontal cortex, BA 45/47; one cluster of 10 voxels, located in the right anterior cingulate; one cluster of 6 voxels located in the left parietal cortex; one cluster of 5, located in the right parietal cortex; one cluster of 5 voxels is located in the anterior right inferior prefrontal cortex, BA 45/47, and one cluster centered around [-3, 39, 3] (see Table 10 above).

Discussion

The fMRI data indicate that both regular and irregular past tense formation use left prefrontal areas: overlapping, but not identical areas in BA 44/45 are active in both conditions, and BA45/47 is active only in the irregular condition. This indicates an anterior involvement for irregular verbs, in contrast to the Words-and-Rules hypothesis. On the other hand, this still constitutes a dissociation between the conditions, in contrast to what the connectionist model predicted. Moreover, since BA 45/47 has been widely implicated in semantic retrieval processes, its activation in the irregular condition should perhaps be not too much of a surprise (see below for a detailed discussion).

As for BA 44/45, no good methodology currently exists to tease apart the two overlapping clusters. Assuming they are identical, this leaves the question of the functionality of Broca's area open. Given the strong evidence that it is active in grammatical processing, it would stand to reason that it would be active in grammatical processing in this task as well. One possibility is that it supplies the feature of 'past-ness' to both conditions early on when the verb is processed in the lexicon, and is hence active in both[12].

In addition, no temporal areas were found to be significantly more active in the irregular condition than in the regular condition. This could be due to a number of reasons. The easiest suspect is the subtraction method; since both stem and past conditions should involve lexical access, subtracting the stem condition from the past condition would result in temporal areas being taken out of the analysis. However, the All-fixation contrast, in addition to the SingleCondition-Fixation contrasts, did not show much temporal activation, either. Thus, the subtraction of trial conditions is probably not the culprit.

Another possibility is that the lack of temporal activation is an artifact due to the experimental paradigm used: during the baseline condition, subjects were given the cue word 'FIXATE", which could have resulted in strong activation of lexical areas for the baseline condition already. A subsequent All-fixation or SingleCondition-fixation subtraction could then show no significantly higher activation for the trial conditions than the baseline condition. This is not so far-fetched, as another language and memory study, run on the same magnet with a similar event-related design and the same subjects, also failed to yield any activation in temporal areas (Dav Clark, p.c.)[13].

---

[12] Supplying the past-ness feature might be separate from concatenating the stem to the affix; if it is, one might expect a 'double-dip' effect for regular past tense formation. This 'double-dip' could not easily be observed without much faster time course data than what fMRI can currently offer, unless it resulted in much higher activation for the regular condition for the irregular.

[13] In general, recording from temporal lobes is generally not robust.

Finally, both conditions use parietal areas which have been implicated in working memory, although, as discussed in chapter 2, the involvement of working memory in the regular past tense formation is implausible (see below for a detailed discussion).

Chapter 6.

General Discussion

Table 13 shows a summary of the data, juxtaposed with the predictions of the two hypotheses and data:

Put Table 13 about here

To repeat, the Words-and-Rules hypothesis predicted a posterior-anterior split between regular and irregular past tense formation, since the irregulars involve a lexical lookup procedure, which should be localized to the temporal areas, whereas the regulars should involve both temporal areas (early stages, lexical lookup of stem) and frontal areas (late stages, application of the rule). In contrast, the connectionist model predicted no neural differences between regular and irregular processing.

From the above table, one can see that the fMRI and MEG data present evidence in favor of the Words-and-Rules theory, since there are some dissociations between the neural computations regular and irregular past tense formation depend on, over the connectionist model, even though they do not match the Words-and-Rules predictions exactly.

The MEG data correlate better with the Words-and-Rules predictions than the fMRI data. As hypothesized, there was a dissociation, both in time and space, between regular and irregular past tense formation: both processes started out in temporal areas, but the regular dipoles jumped anterior, possibly to Broca's area, while the irregular dipoles remained in temporal areas. Moreover, the regular past condition had a faster time course than the irregular past condition, as predicted by the reaction time data. In the stem conditions, no significant difference, both in time and space, could be seen between regular and irregular verbs[14].

However, the fMRI data are more difficult to reconcile with the stated Words-and-Rules hypothesis, that there should be an anterior-posterior dissociation between regular and irregular past tense formation. While the fMRI data confirm a frontal involvement for regular verbs, they also indicate a frontal involvement for irregular verbs as well. Furthermore, there was no activation in temporal areas for either regular or irregular verbs, as discussed above.

There are two main possible interpretations of the cluster in the anterior left inferior cortex, BA 45/47, which was significantly more active in the irregular subtraction condition than in the regular subtraction condition, which involve a modification of the original hypothesis. First, this cluster could be involved in the *inhibition* (or control) of the default rule in the irregular past condition. Second, it could be responsible for semantic retrieval and selection of the correct irregular form.

---

[14] This is assuming that the dipole localizations can be trusted.

With respect to the first option, many imaging experiments have implicated left prefrontal areas in the control of cognitive processes (e.g. in suppressing the pre-potent response, which is reading the word in the color it is written in, instead of the word itself, in a modified Stroop test (MacDonald III et al., 2000); or in suppressing the pre-potent response in Piaget's A-not-B task, which is looking for the object in the old location (A) instead of the new (B) (Diamond et al., 1989). However, a inhibitory role for BA 45/47 in the production of irregular past tense forms seems to be unlikely, because the left prefrontal area most often mentioned as the locus of control is Dorsolateral Prefrontal Cortex (DLPFC) or BA 9/46, and not BA 45/47 (see e.g. Miller and Cohen, in press; Goldman-Rakic, 1987; Bunge et al., 2000; Fuster, 1997, for a discussion of prefrontal involvement of cognitive control).

The second possibility, that BA 45/47 is involved in the selection of the correct irregular form, is more plausible. There is extensive brain imaging literature that indicates that BA45/47 is involved in semantic and/or lexical retrieval and selection during language tasks, with two possibilities: either BA 45/47 is engaged in semantic retrieval *per se*, i.e. it facilitates access to relevant semantic knowledge, or it is necessary for the *selection* of task-relevant representations from among competitors.

In favor of the first possibility, Desmond et al. (1998) showed this area was active in a stem-completion paradigm, which clearly involves selection from the lexicon; Petersen et al. (1988) showed that BA 45/47 is selectively active during the recall of a semantically related word to a visually presented stimulus (e.g. *eat* to *cake*). Furthermore, Wagner et al. (1998) and Kirchhoff et al. (2000)

implicated BA 45/47 in the process of correctly encoding verbal stimuli (words) as opposed to non-verbal stimuli such as pictures, and Poldrack et al. (1998) showed that semantic processes (abstract/concrete decision) activated BA 45/47 more than phonological ones (syllable counting). In addition, Demb et al. (1995) showed that while BA 45/47 was more active in the semantic (abstract/concrete decision) than in the non-semantic (upper case/lower case or ascending/descending judgements) condition; they also showed that this differential activation was not due to the semantic task being harder, since BA 45/47 was more active in the semantic condition than in the ascending/descending judgement, even though the latter task was more difficult (as shown by longer RTs for the latter task.).

In favor of the second possibility, Wagner et al. (submitted) suggested that BA 45/47 is engaged in controlled semantic retrieval, i.e. BA 45/47 may be active under retrieval conditions in which there are no pre-potent responses. They used a similar paradigm to Thompson-Schill et al.'s (Thompson-Schill et al., 1997, 1998, 1999), in which a target word had to be selected from a list given a cue word, but kept the selection demands to a minimum, and showed that BA 45/47 was reliably active only in conditions in which the cue-target association strength was weak, or when the size of the target group increased (see Wagner (in press) for an overview of the literature).

To apply these findings to the current study, irregular past tense forms are stored in memory with some associative properties. It is plausible that while *brought* has to be retrieved as the correct form for *bring*, the associative properties of the memory stores could also bring incorrect forms on line, which could be correct, by analogy to other irregular verbs. In the case of *bring*, the incorrect forms brought on line would include *brang* (by analogy to *sing*) and *brung* (by analogy to *fling*). BA 45/47 could be involved in selecting the correct form from the possible choices, even though there is a prepotent response (the correct form).

This makes a hypothetically verifiable prediction: irregular verbs which reside in dense phonological neighborhoods, such as *bring, sing,* and *fling*, should activate BA 45/47 more strongly than irregular verbs which do not, such as *eat* and *build*. The obstacle to this verification would the small number of verbs available for the study, which would greatly reduce the signal to noise ratio.

Alternatively, BA 45/47 could be involved in simply retrieving the form that corresponds to *bring* + past feature; in this model, BA 45/47 would just retrieve the correct form brought, without ever having to bring other forms such as *brang* or *brung* on line.

The two overlapping clusters, which correspond to left BA 44, or Broca's area, are more difficult to interpret. There are two main possibilities, depending on whether they represent two distinct clusters or not. Presently, it is impossible to reach a conclusion, given the data; therefore, for the purpose of discussion, I will assume these clusters are not distinct. Since both irregular and regular past tense formation make use of this area, one possibility is that it supplies the grammatical feature of 'pastness' or 'finite-ness', or any other feature that distinguishes past tense forms from stems, to the overall process of past tense formation, in addition to concatenating the stem with the suffix, as discussed above. Another is that this area is in fact involved in appending the suffix *–ed* to the stem in the regular condition, and it is active in the irregular condition as well, since the regular rule applies by default. For irregular verbs, the incorrect, overregularized form (such as *bringed*) would have to be selected against. Again, this selection could the be role of BA 45/47. This would be a slight modification to the Words and Rules theory, in that the default rule no longer has to be blocked from applying to irregular verbs; instead, the rules applies anyway, and only the incorrect, overregularized form has to be selected against.

The parietal regions which show a significant activation for the double subtraction conditions are similar to those implicated by Jonides et al. (1998) and Awh et al. (1996) in verbal working memory paradigms, in which subjects were asked to memorize and retrieve novel words. In these studies, posterior parietal regions, BA 7 and BA 40, were reliably active during both the storage and the retrieval conditions, but not during the encoding condition. While it might be plausible to posit a semantic retrieval component for irregular verbs, however, a short term storage component is difficult to reconcile with both the speed and automated nature of regular past tense processing, or even the retrieval process of irregular forms.

Finally, the anterior cingulate, which is active only in the irregular double subtraction condition, is known to co-vary frequently with left frontal activity, although the manner in which this happens is still unclear. It has been implicated in an array of related function, in making and monitoring of decisions (Liddle et al., 2000), in controlling or inhibiting a prepotent response (MacDonald II et al., 2000; Rubia et al., 2001), attentional regulation and feedback (Thayer and Land, 2000), and monitoring on-line processing (Ochsner et al., 2001), to name a few. How any of the above functions could have a bearing on the formation of the English past tense, if any, is a subject of further investigation.

Another open question is why the MEG data do not indicate any frontal involvement for the formation of irregular past tense forms; this was possibly due to the ECD analysis looking to localize only 1 dipole, as opposed to several. It is doubtful that the differences between the MEG and fMRI data can all be ascribed to one involving overt (MEG) and the other covert (fMRI) production of words. As reviewed by Fiez and Petersen (1998), the major differences in activation between overt silent language production tasks were more activation in the motor areas (BA4 and 6) and auditory cortex (BA 22), due to activation through one's own voice. The bigger difference is probably the fact that the MEG experiment was run in the past tense production conditions and the stem reading conditions in one block each, whereas in the fMRI experiment, both conditions were randomly intermixed. In addition, the MEG data are not subtraction data, and show the timecourse of activation, which allowed the dissociation between regular and irregular past tense formation in the later stages to emerge.

The combined MEG and fMRI data lend support to the Words-and-Rules theory over connectionist models, in that irregular and regular past tense formation can be dissociated, both in their time course and their neural substrates, in contrast to what the connectionist model predicted. The irregular past tense formation activates temporal areas, and the regular past tense formation activates frontal areas. A modification is suggested, in light of the fMRI data, which indicate frontal involvement for irregular forms as well. The details, however, still have to be worked out; it is unclear whether the left frontal areas around Broca's area are distinct clusters or not, and what the role of the anterior cingulate is, or the exact role of BA 45/47, why there was no temporal-lobe activation overall in the fMRI data, and whether the parietal areas found in the fMRI data are indeed connected to verbal working memory and storage. More experiments will have to be conducted in order to answer these questions.

Future experiments would include the following: a re-analysis of the current fMRI data without smoothing; this will reduce the spatial extent of the data, and thus perhaps separate the two overlapping clusters located in BA 44/45. For the same purpose, the same fMRI experiment could be run, either on a stronger magnet, or with surface coils focused on left prefrontal areas instead of a whole head coil, in order to increase the signal to noise ratio.

Another possible re-analysis would involve looking at irregular verbs in high phonological density neighborhoods vs. low density neighborhoods. If BA 45/47 is indeed involved in selection, verbs in low density neighborhoods should show less activation in BA 45/47 than verbs in high density neighborhoods. For this experiment, more subjects would have to be recruited, since dividing the stimuli into smaller bins would drastically reduce the signal to noise ratio of the currently available data set. Alternatively, one could present the same set of verbs multiple types per subject.

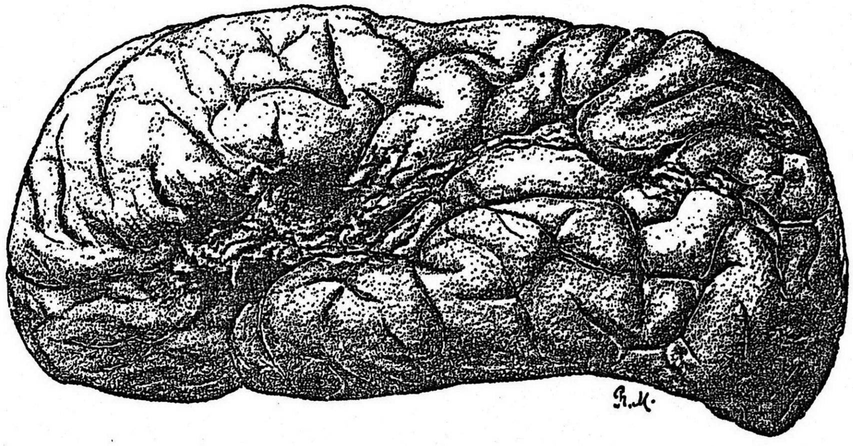
Another possible experiment would look at the putative working memory components of the past tense formation, by making the process so hard that it would become less automated and more dependent on working memory, perhaps by embedding a few regular verbs in many irregular verbs and vice versa.

Finally, in order to correctly merge the MEG with the fMRI data sets, the MEG experiment would have to be re-run with exactly the same event-related design as the fMRI experiment.

In sum, the work presented here indicates that words and rules can indeed be dissociated in the brain, and that further work, using both techniques, should shed further light on the intricacies of neurolinguistics.

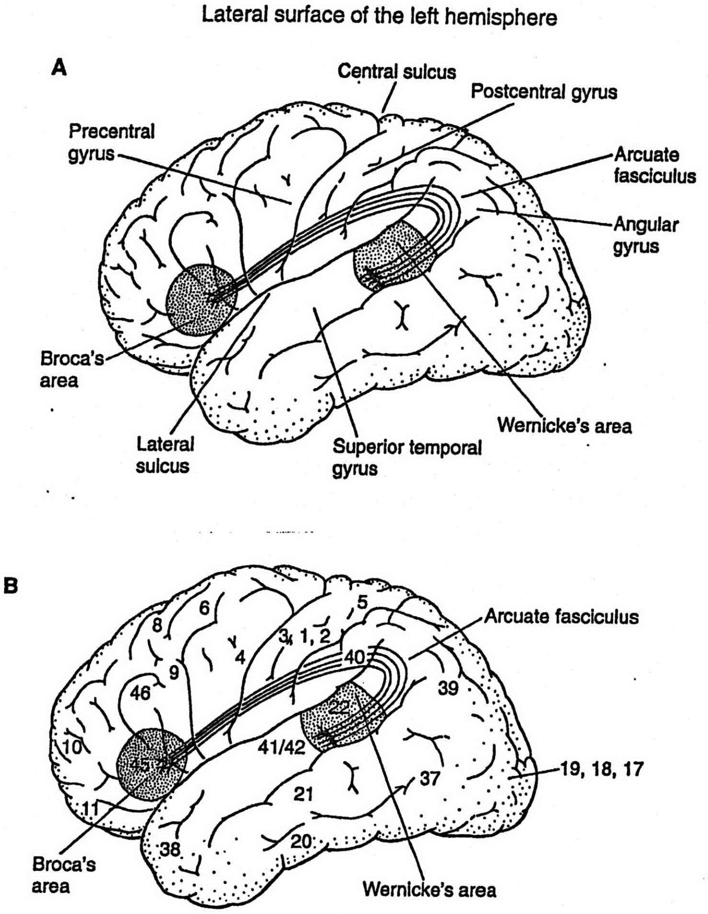
Figures

Figure 1
Broca's area



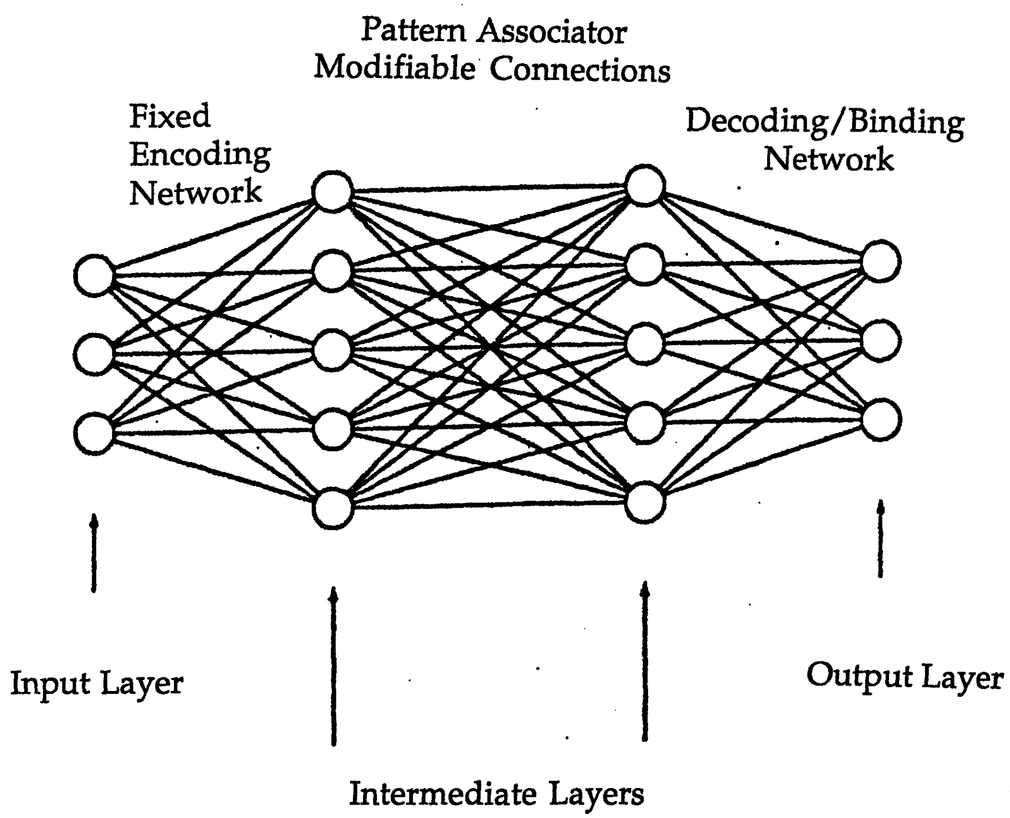
The brain of Lebourgne showing a lesion in Broca's area (*Source*: Moutier 1908: 78)

Figure 2
Broca's and Wenicke's areas



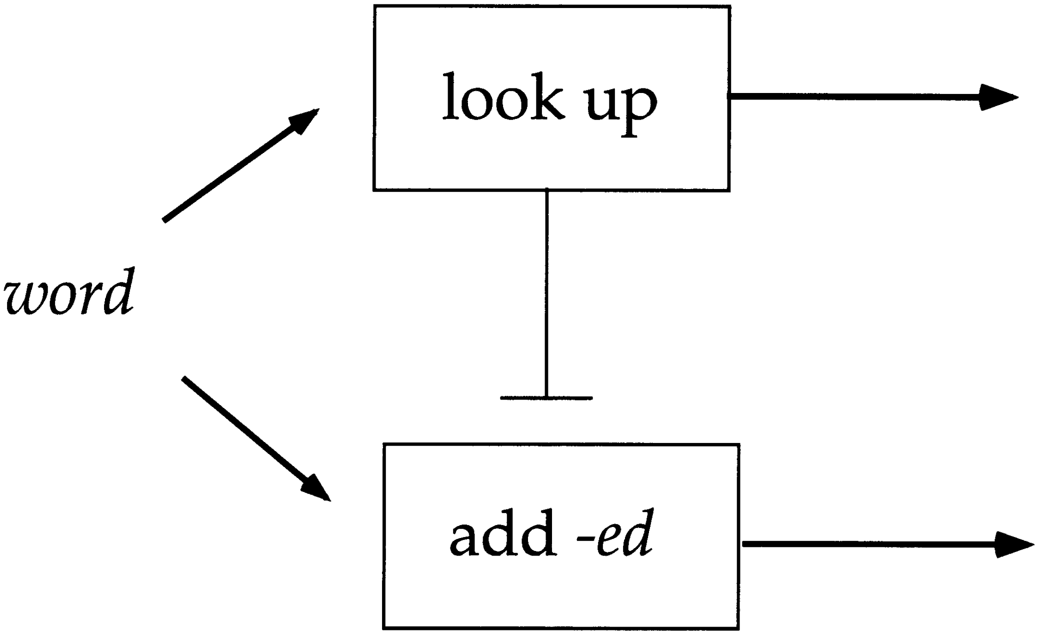
Source: Kandell et al.,1991

Figure 3
Basic Neural Network



(Adapted from: Prince and Pinker, 1988)

Figure 4
The Words and Rules Past Tense Model



Source: Pinker, 1999

Figure 5
The Words and Rules Model in the Brain

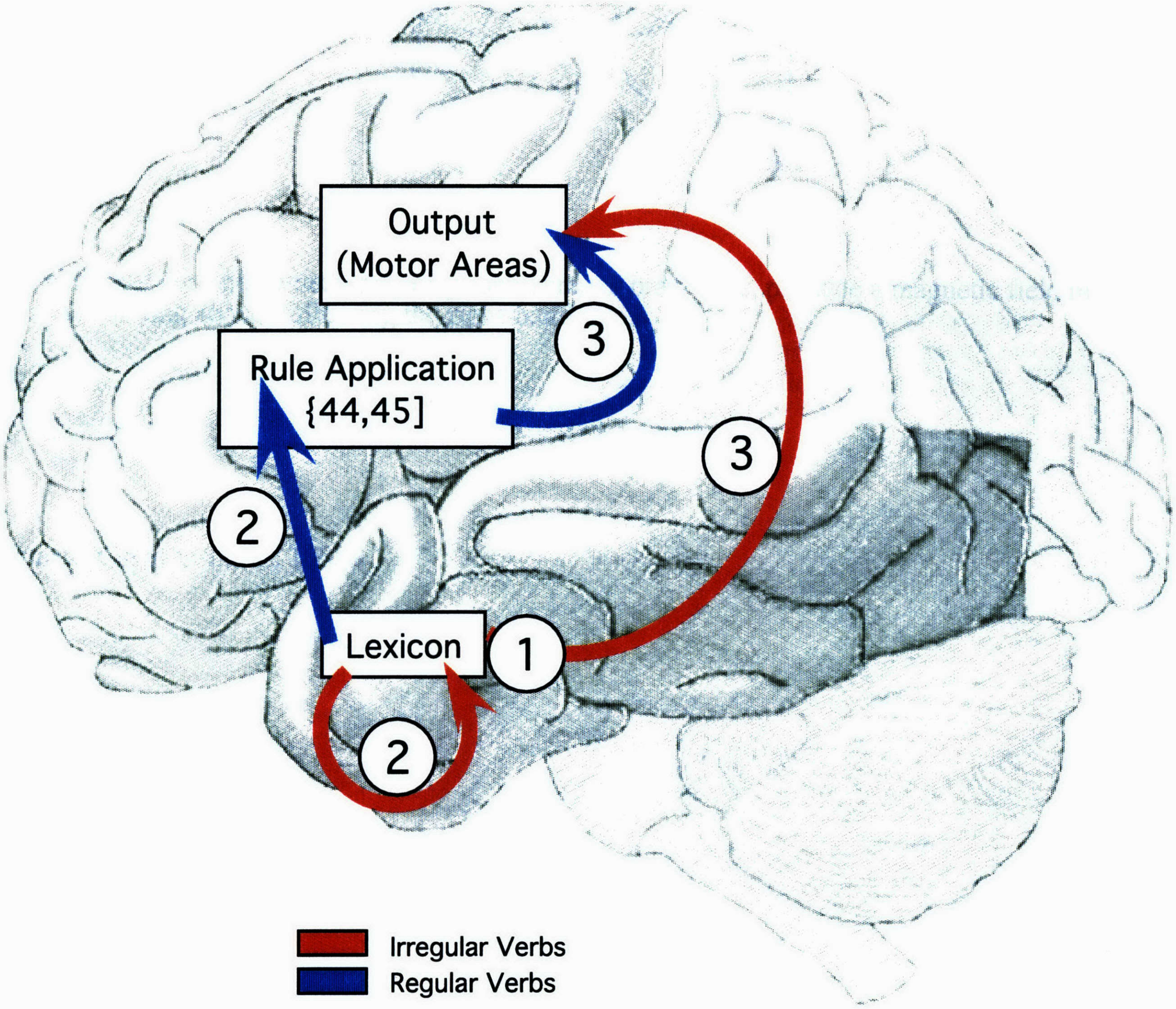
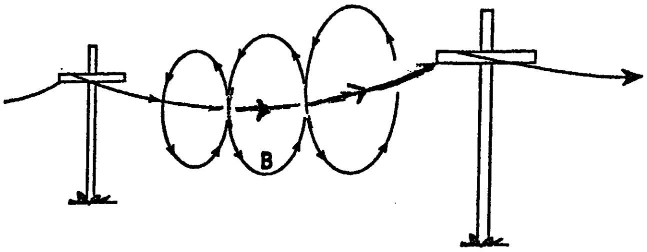


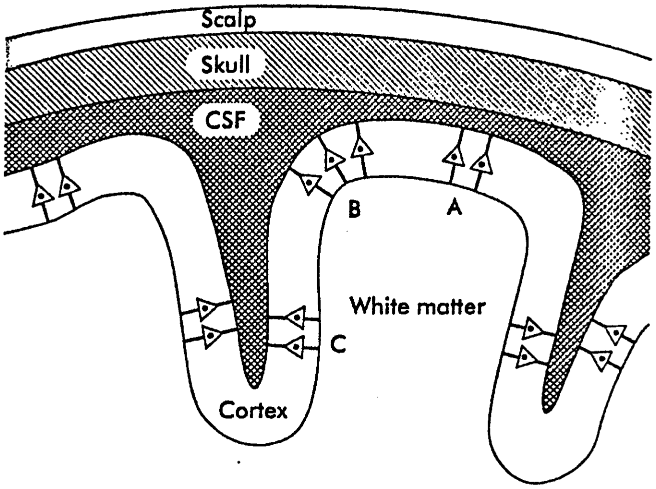
Figure 6
Magnetic Field Lines Around Current Flow



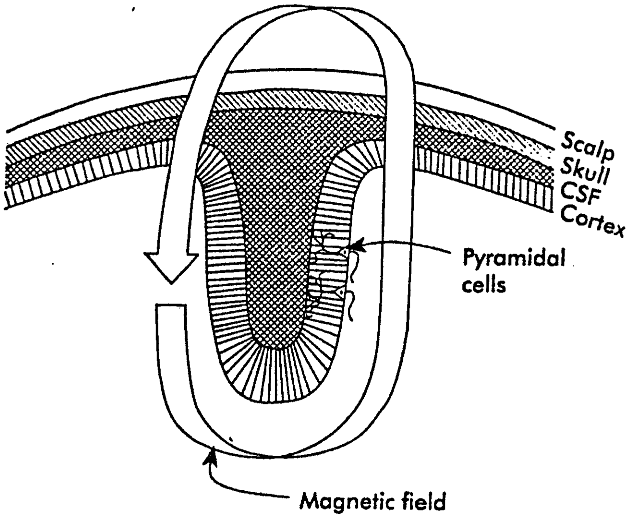
All electrical currents, whether in telephone wires or brain cells, produce a magnetic field in the surrounding space, following the right-hand rule.

Source: Orrison et al., 1988a

Figure 7
Cortical Cells and Their Contribution to the MEG Signal

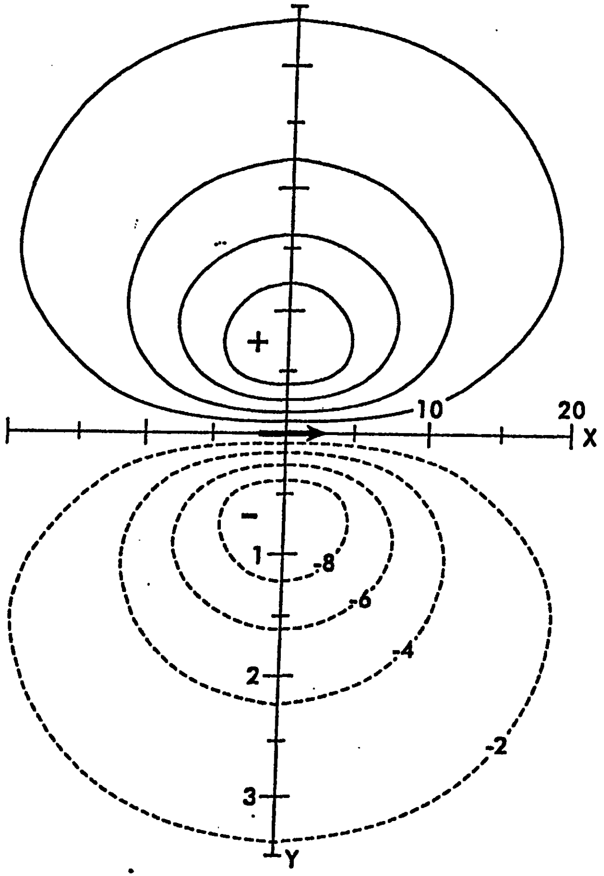


Cells oriented perpendicular to the skull surface (A) fail to generate an extracranial magnetic field. Cells oriented parallel to the skull surface (C) produce a significant radial magnetic field which can be picked up the MEG sensors. Cells of intermediate orientation (B) have both radial tangetial components.



Source: Orrison et al., 1988a

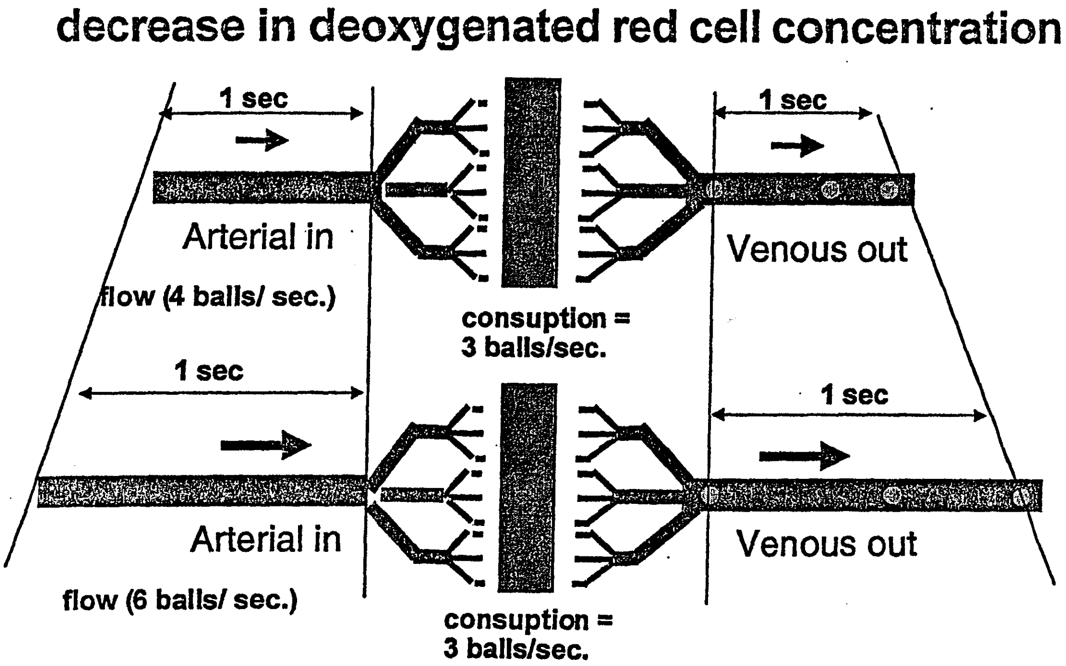
Figure 8
A Current Dipole



Isofield contour map of the magnetic field generated by a dipole embedded in a half-space below the plane of measurement. The dipole (*arrow*) is located below the origin and oriented along the x axis.

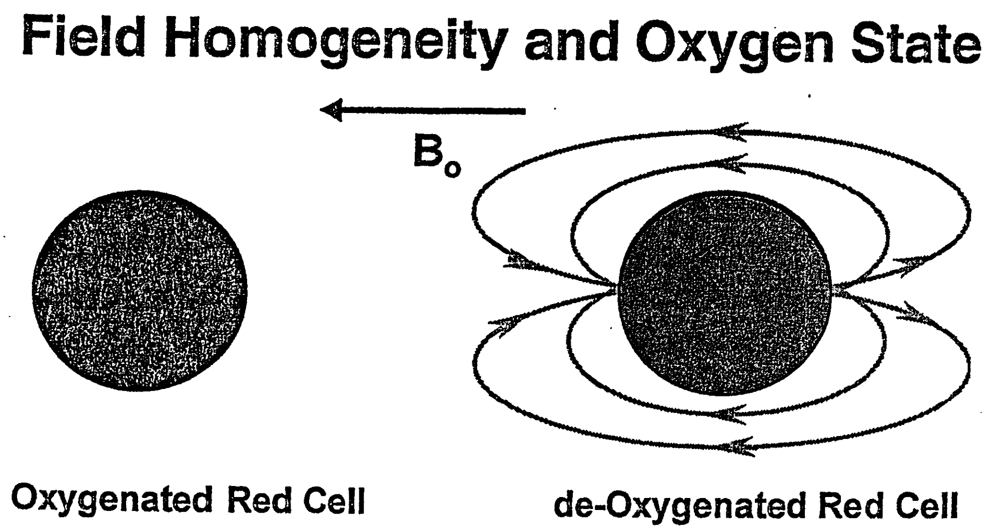
Source: Orrison et al., 1988a

Figure 9
Model of Increased Blood Flow in Active Brain Areas



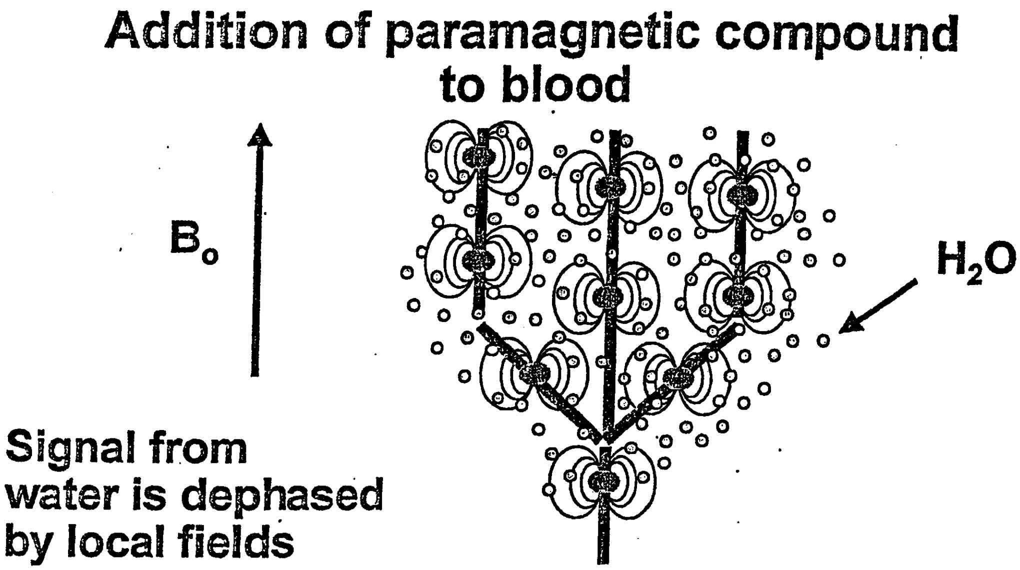
Source: Visiting Fellowship in fMRI, 2000

Figure 10
Oxygenated vs. De-oxygenated Blood



Source: Visiting Fellowship in fMRI, 2000

Figure 11
The BOLD Effect

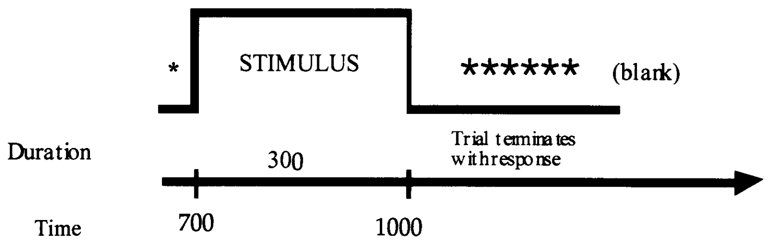


B_o = applied magnetic field

Source: Visiting Fellowship in fMRI, 2000

Figure 12
Scan Sequence of the MEG Experiment

TRIAL STRUCTURE



SCAN TYPES

Regular Past *Irregular Past*	*	WALK EAT	******
Regular Stem *Irregular Stem*	*	TALK SWIM	******

Figure 13
A Sample MEG Isocontour Map

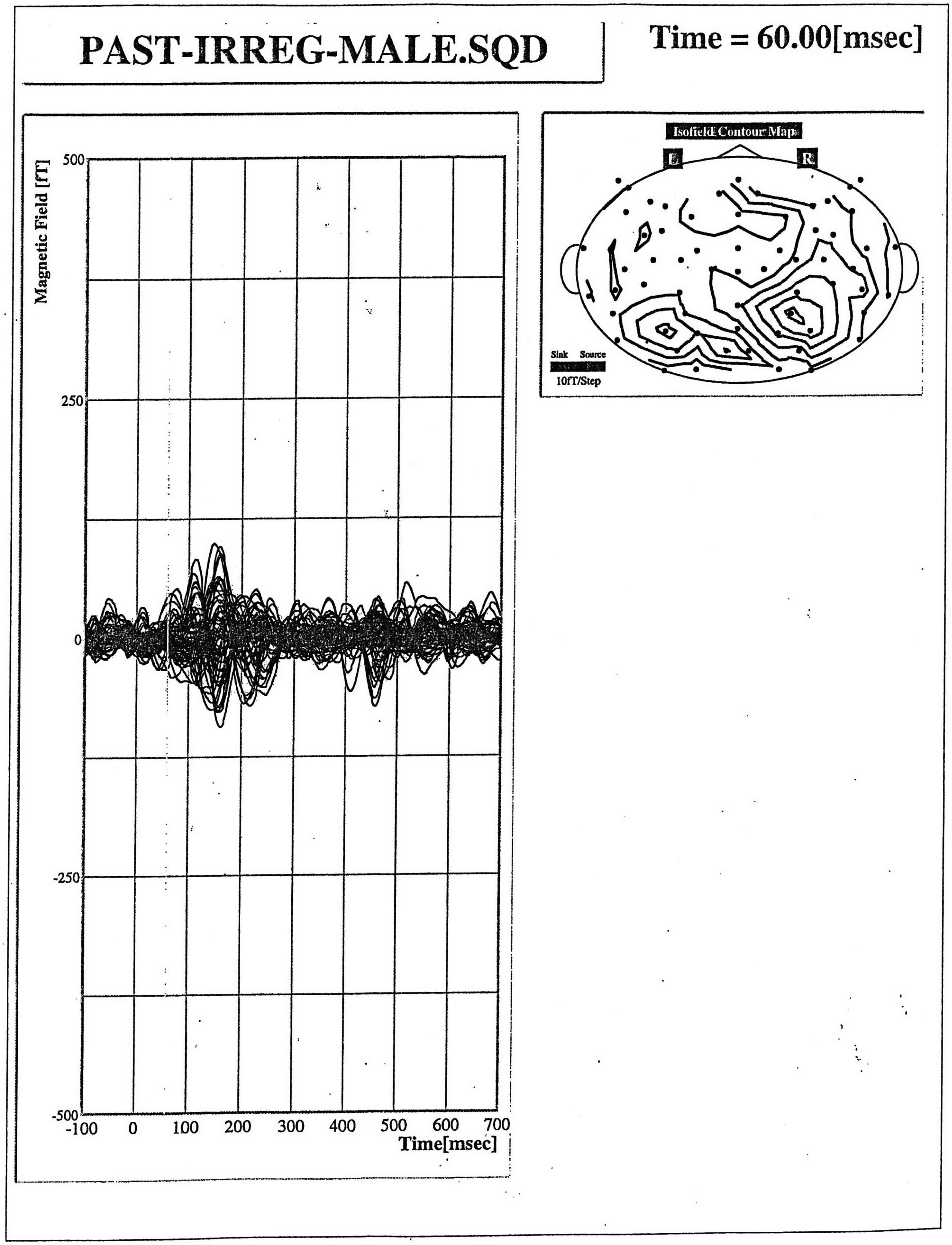


Figure 14
A Sample RMS graph

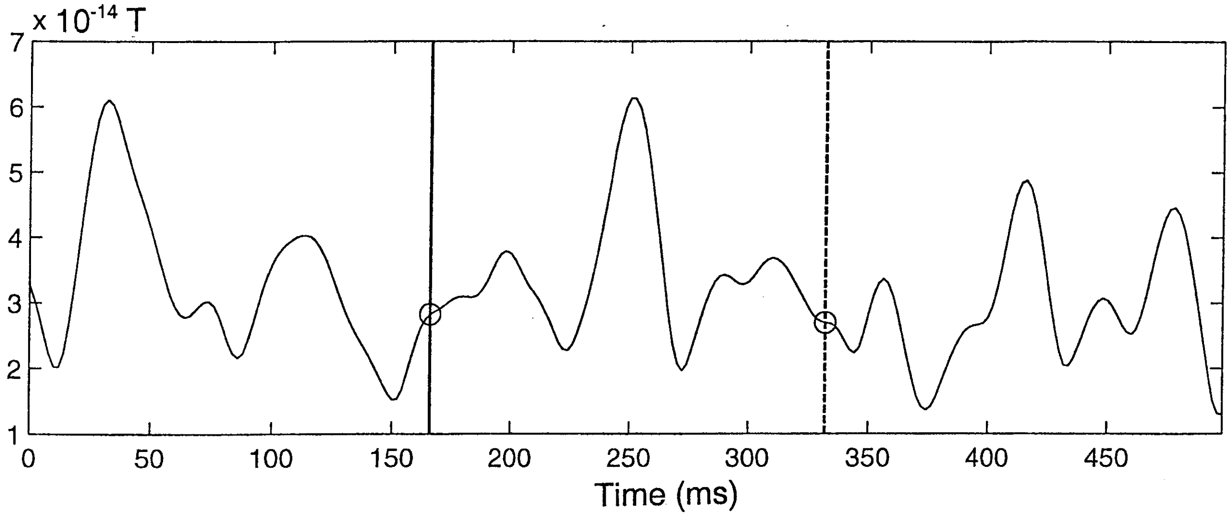
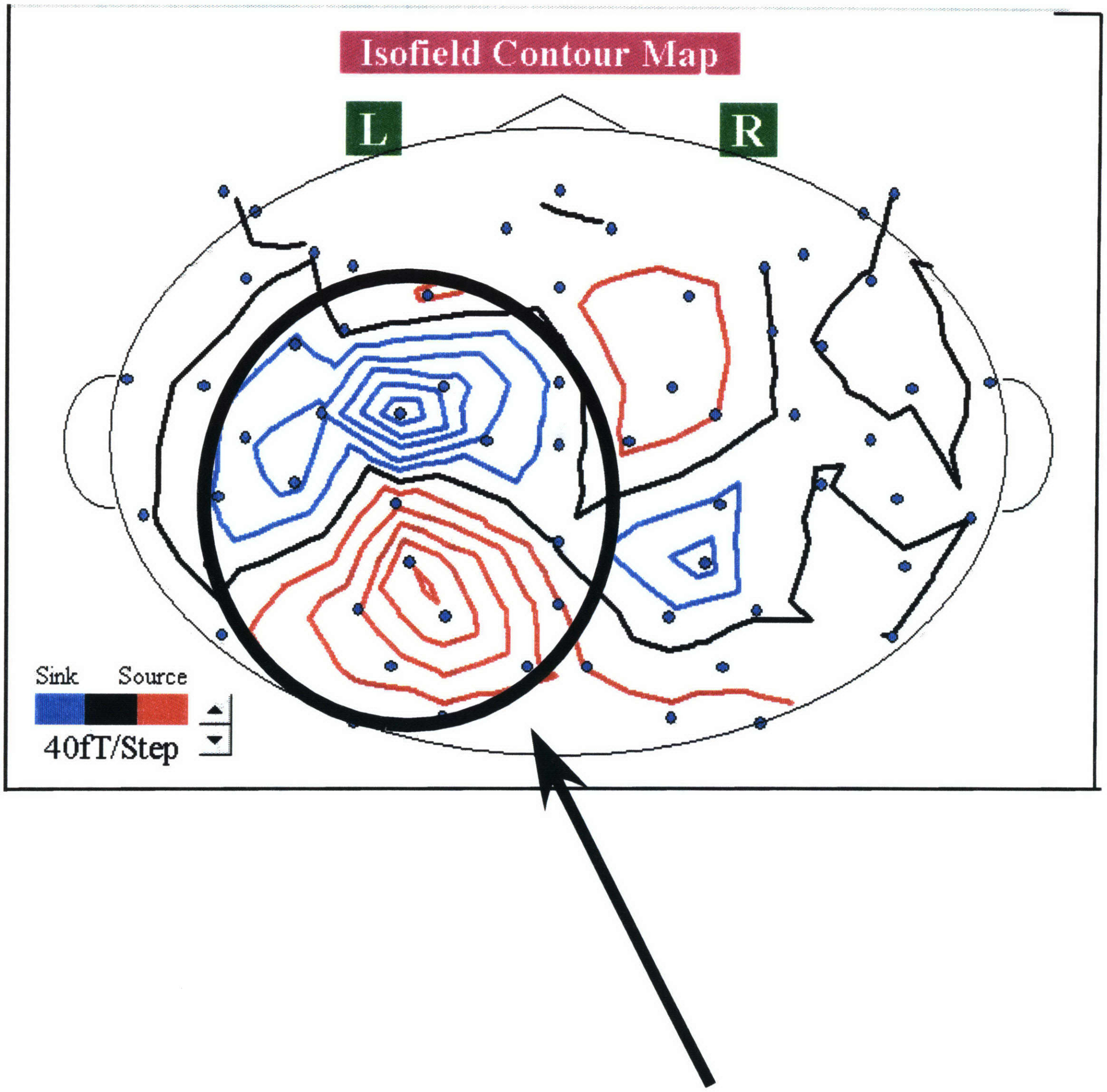


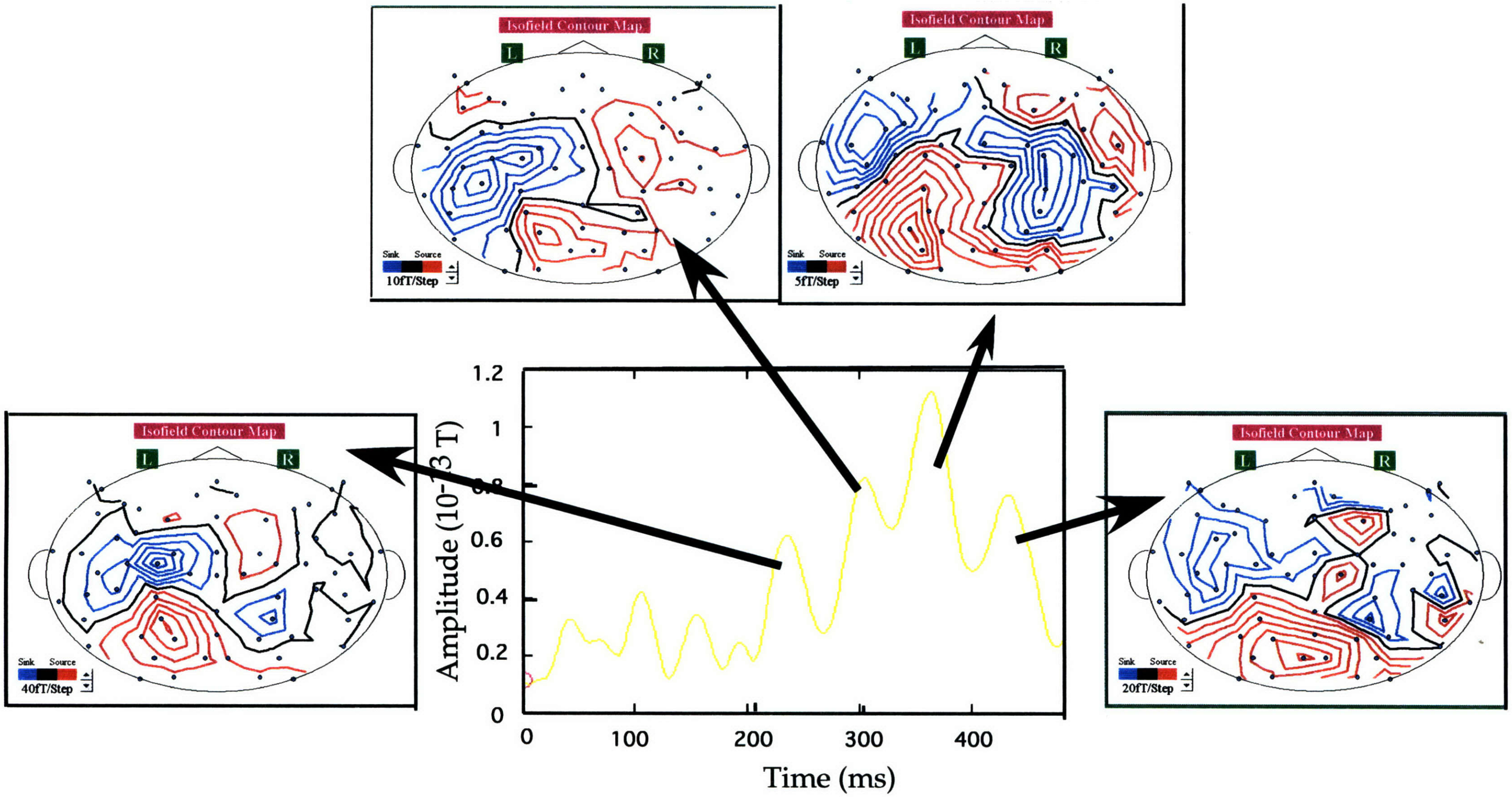
Figure 15
Selection of Channels for RMS Data Anslysis



Channels selected

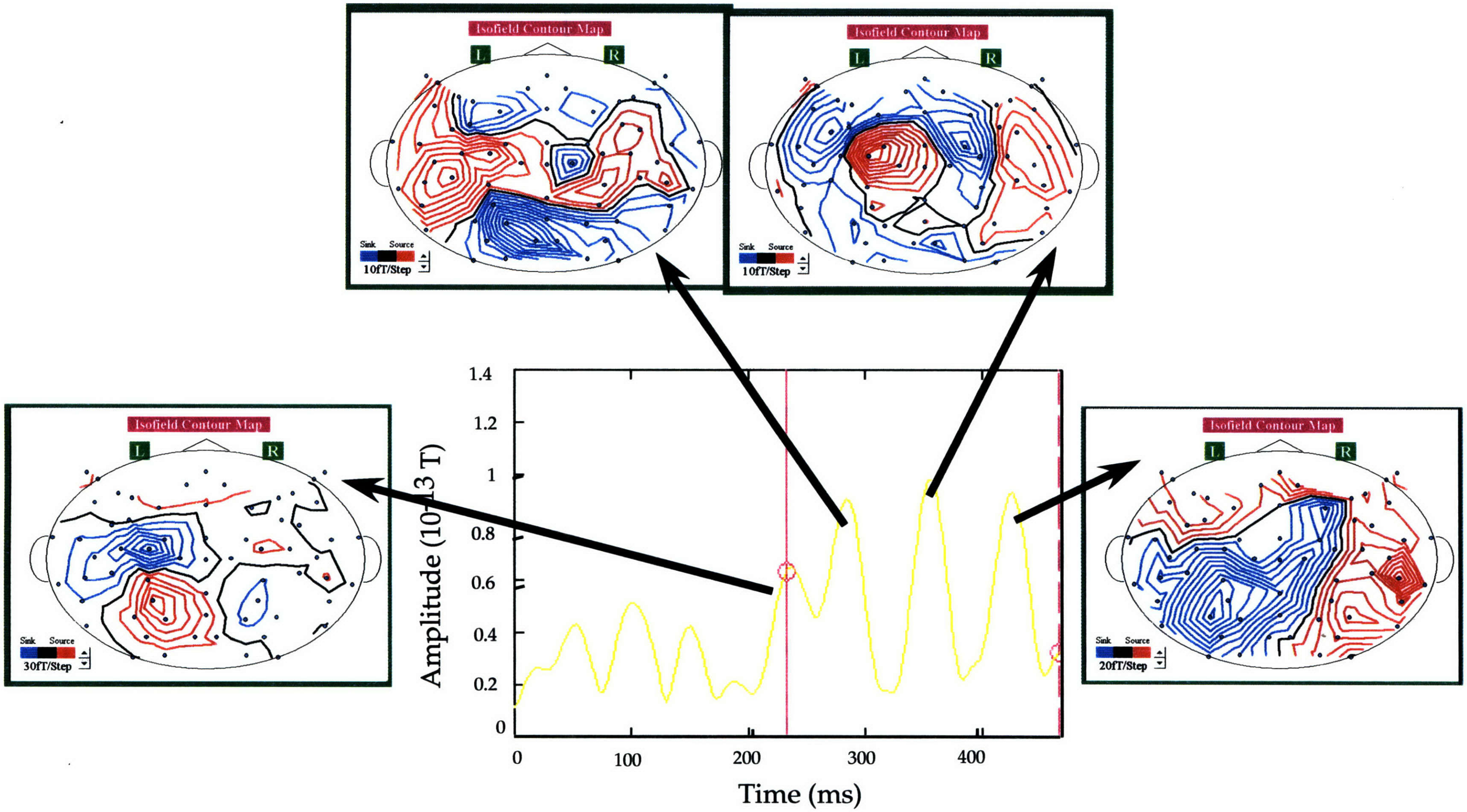
Figure 16
Isocontour Maps For the Four Conditions in the MEG Experiment

Irregular Past Tense Formation



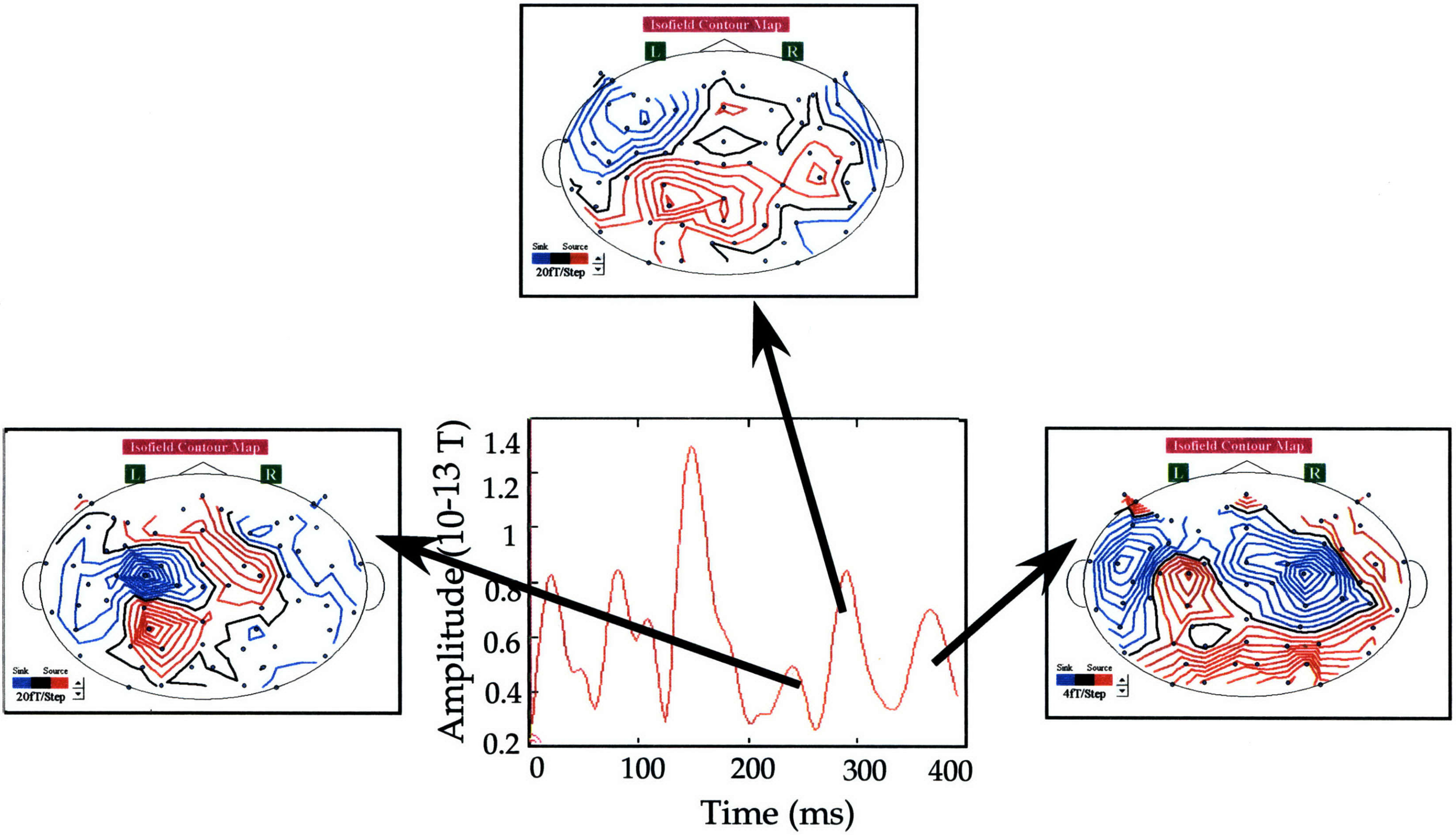
Isocontour Maps For the Four Conditions in the MEG Experiment

Regular Past Tense Formation



Isocontour Maps For the Four Conditions in the MEG Experiment

Irregular Stem Reading



Isocontour Maps For the Four Conditions in the MEG Experiment

Regular Stem Reading

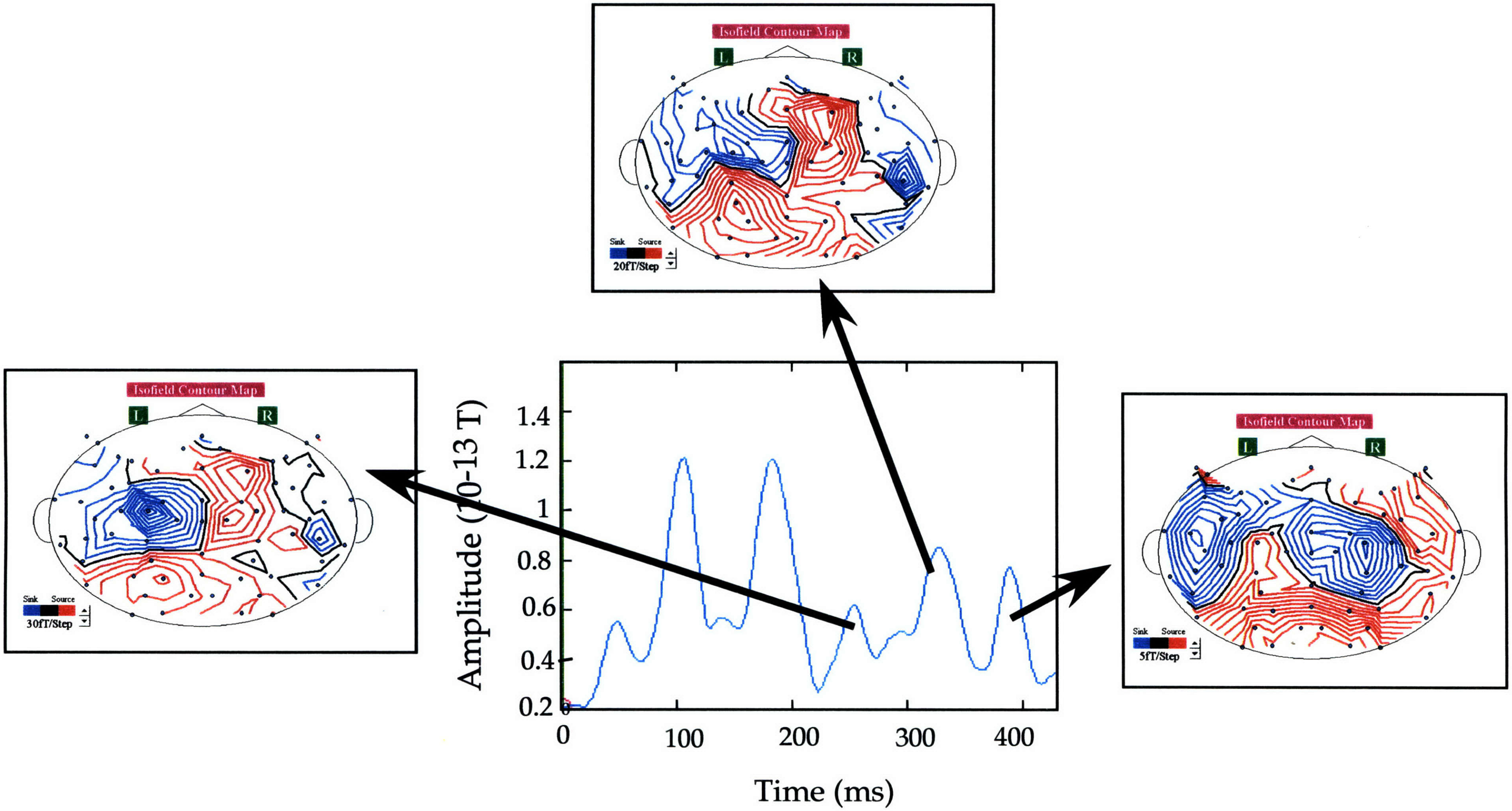
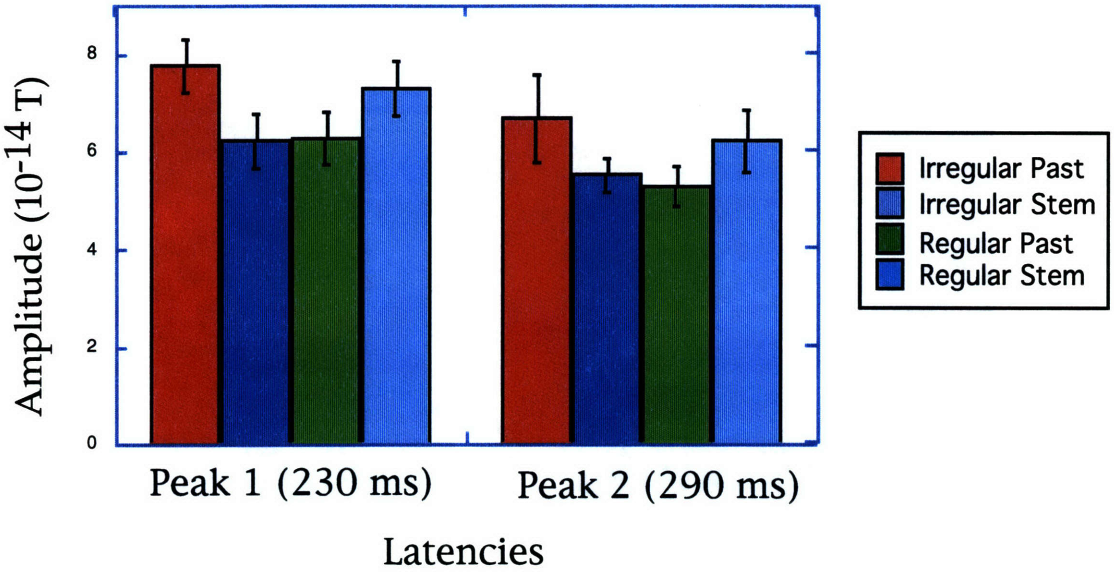
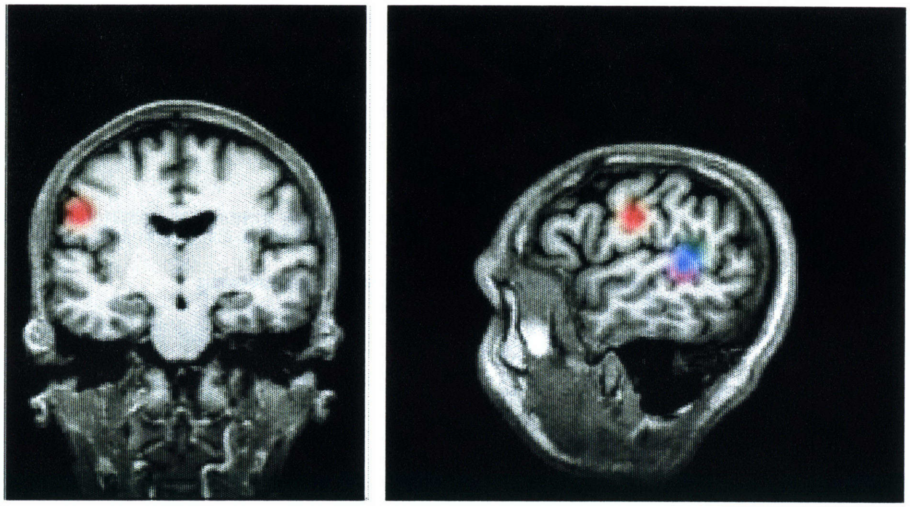


Figure 17
RMS Amplitude Data For the First Two Peaks



These are the amplitude data for the peaks in which the Regularity x Task interaction term was significant at p<.001.

Figure 18
Approximate Dipole Localizations for the MEG Experiment

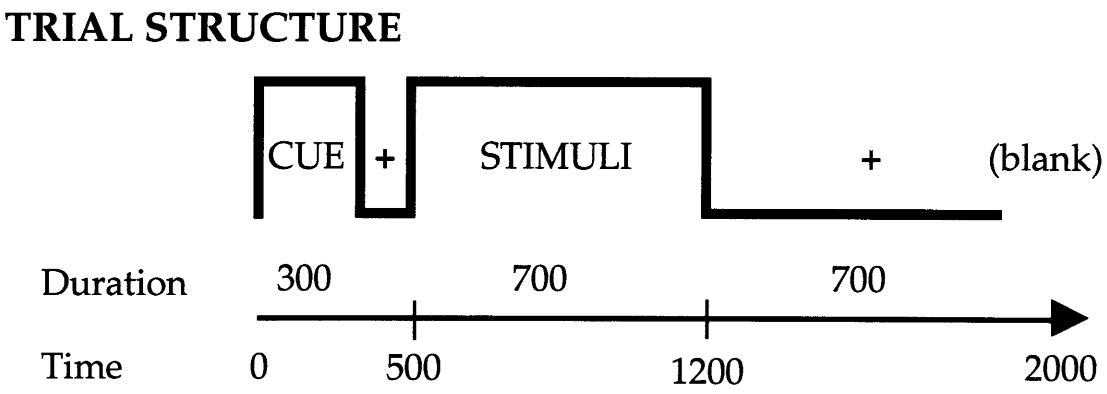


🔴 Regular Past Tense, 350-450 m, [-50, 6, 22], left frontal lobe

🟢 Regular Past Tense, 210-310 ms, [-40, -19, 4], left temporal lobe

🔵 Irregular Past Tense, 210-470 ms,[-40, -20, -10], left temporal lobe

Figure 19
Scan Sequence for the fMRI Experiment



TRIAL TYPES

Fix	FIXATE	+	+
Regular Past *Irregular Past*	PAST	WALK EAT	+
Regular Stem *Irregular Stem*	READ	TALK SWIM	+

Figure 20
All-Fixation Subtraction Data

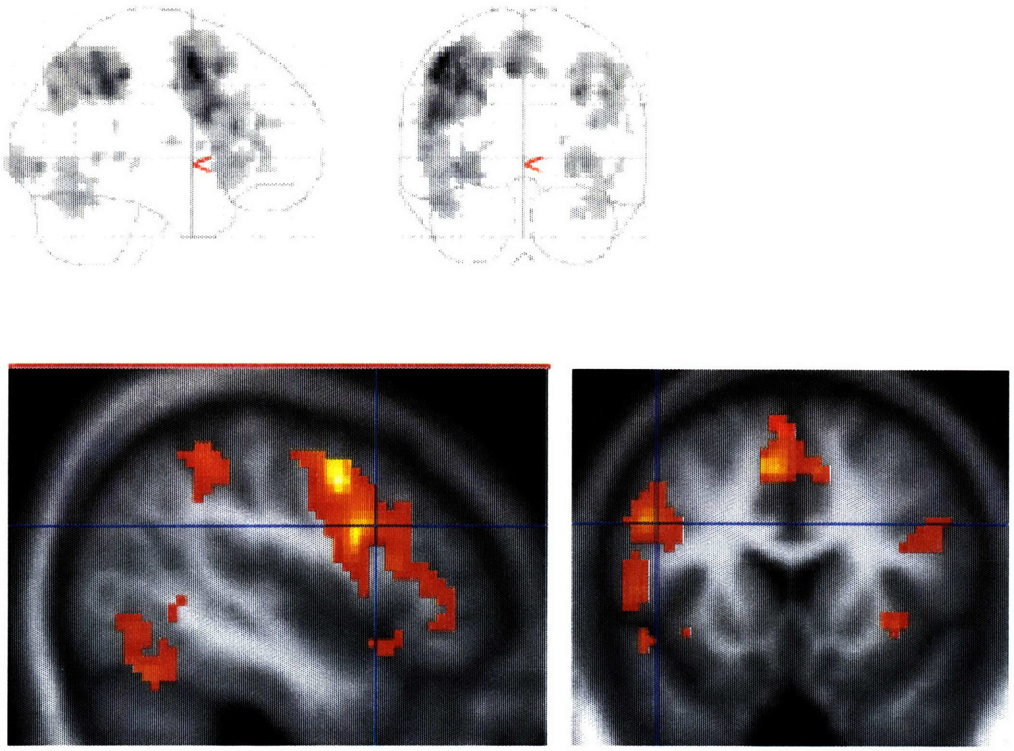


Figure 21
Past-Stem Subtraction Data

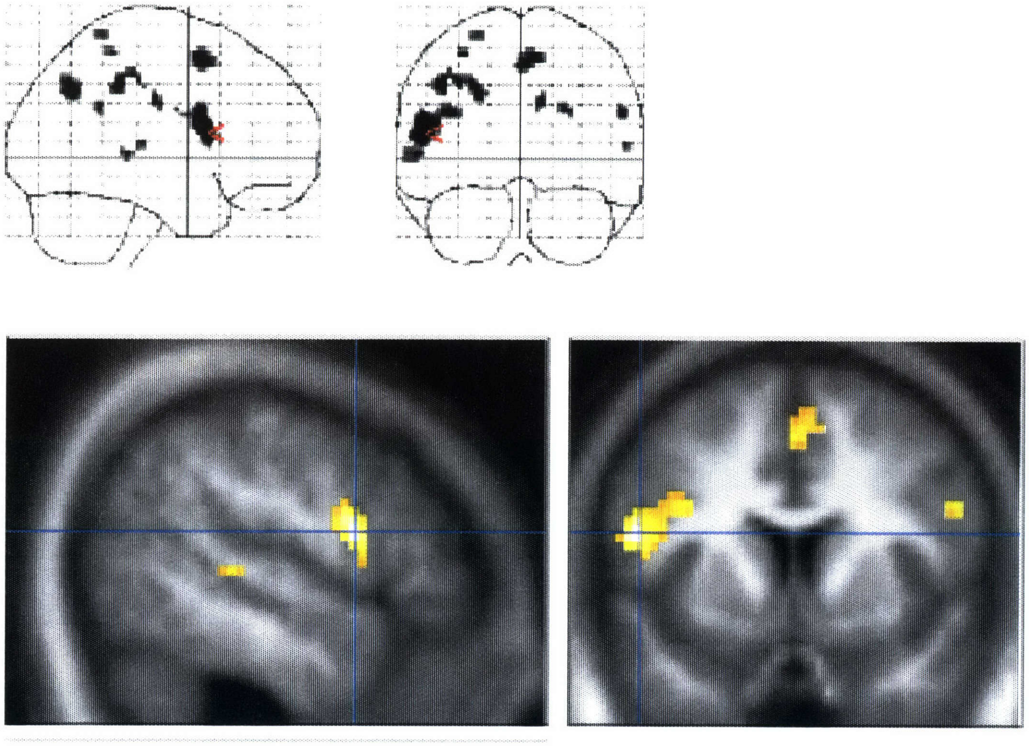


Figure 22.
RegularPast-RegularStem Subtraction Data

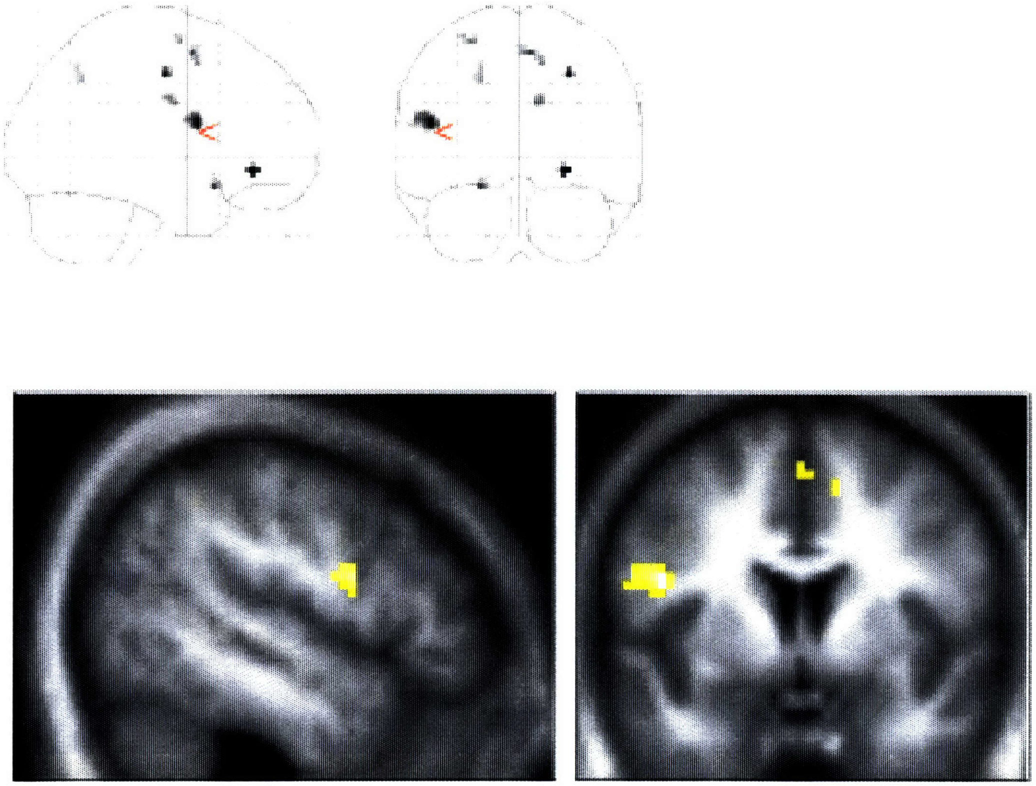


Figure 23
IrregularPast-IrregularStem Subtraction Data

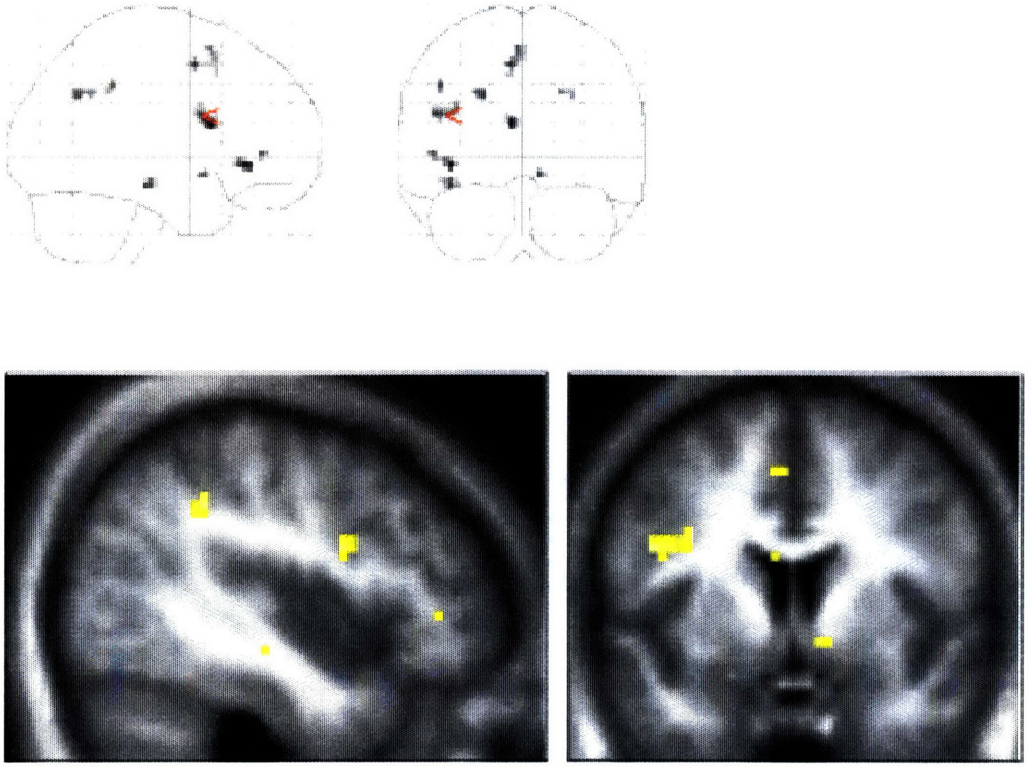


Figure 24

Extent of Overlap Between Regular and Irregular Past Tense Formation in BA 44/45

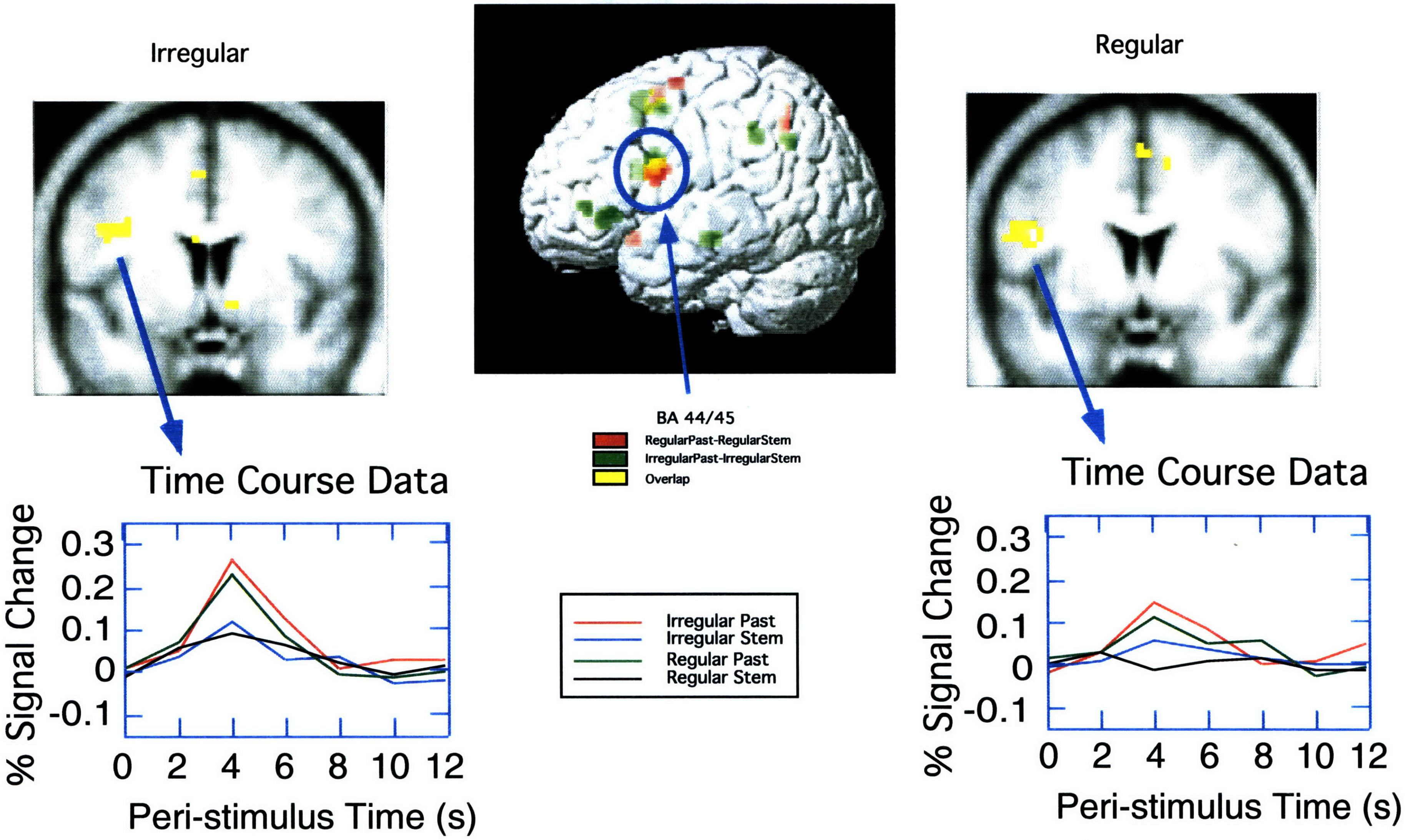
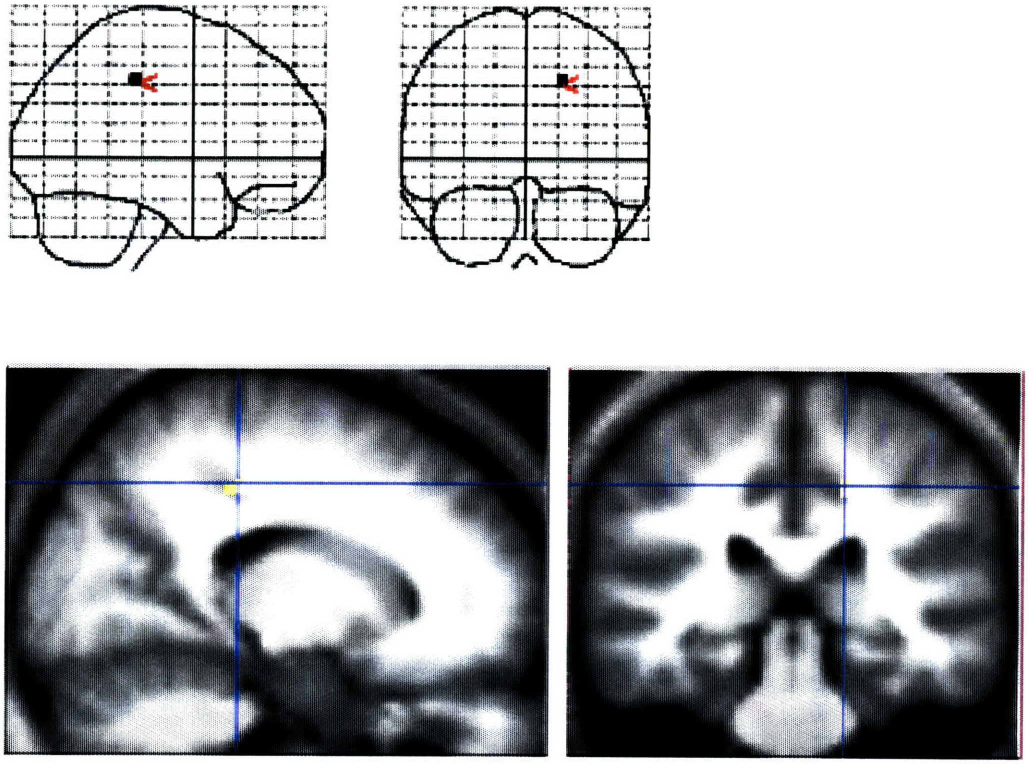


Figure 25
(RegularPast-RegularStem)-(IrregularPast-IrregularStem) Subtraction Data



Time Course Data

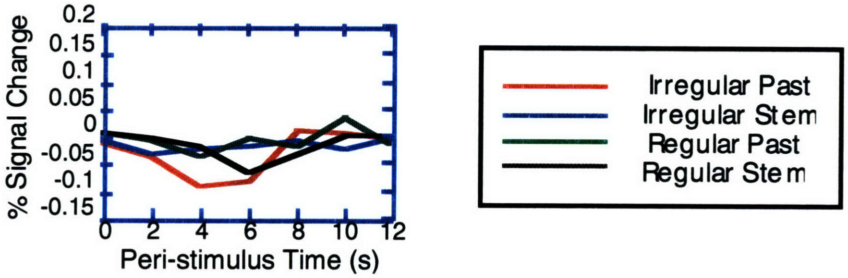
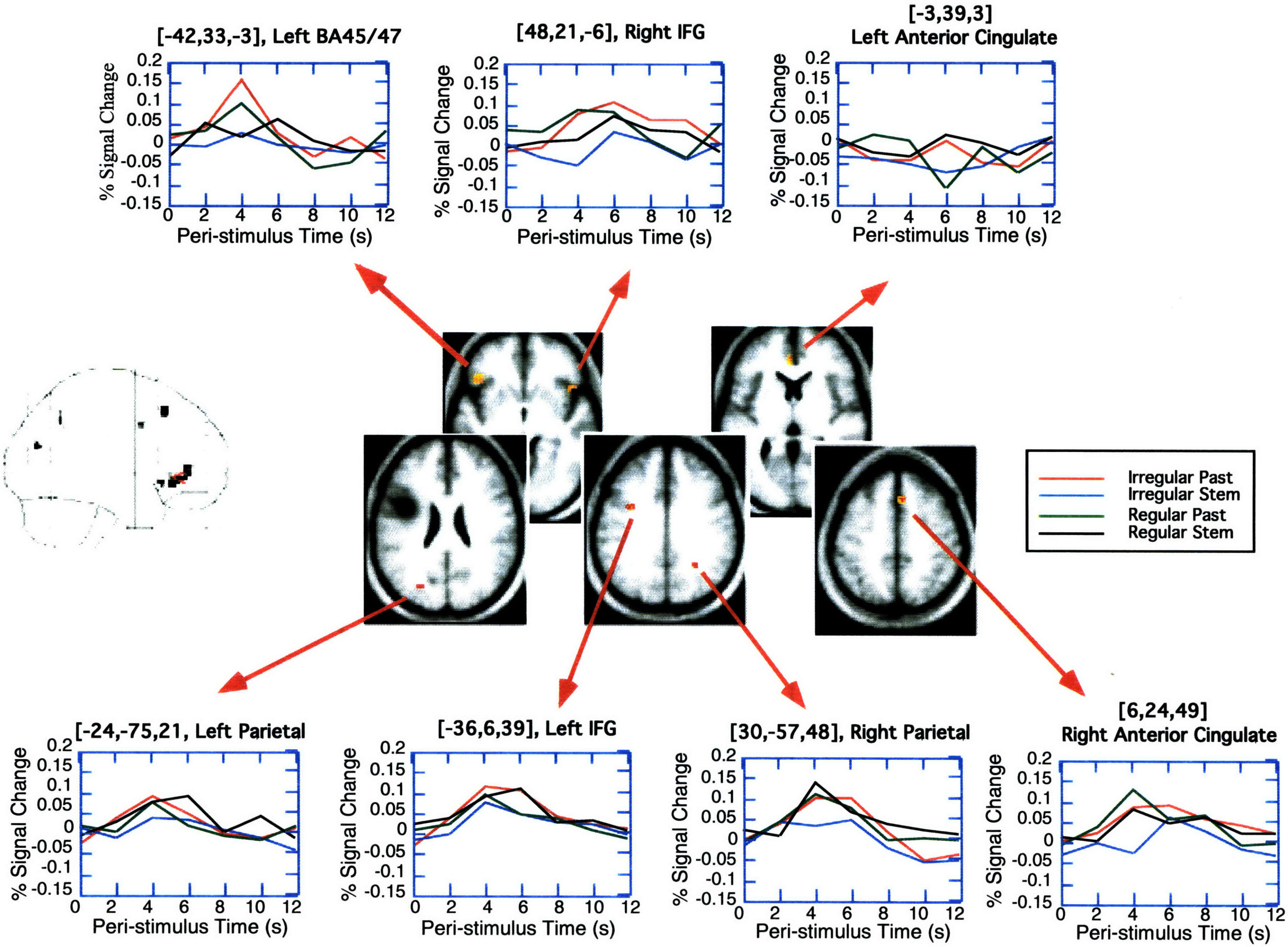


Figure 26
(IrregularPast-IrregularStem)-(RegularPast-RegularStem) Subtraction Data



Tables

Table 1
Regular and Irregular Verbs

	Irregulars *(dig-dug)*	**Regulars** *(look-looked)*
Are matched in:		
Complexity	One phonological word	
Syntax	Tensed	
Meaning	Past	
But are different in:		
Predictability	Very Low (*sing-sang, bring-brought, fling-flung*)	Very High (verb + -/*d*/)
Productivity	Very low (nearly fixed list:180 verbs)	Very High (*faxed, snarfed;* novel verbs: *plagged*)

Table 2
Common Magnetic Sources

Magnetic Source	Strength (fT)
Field applied in MRI	10^{15}
Field near a small bar magnet	10^{13}
Earth's magnetic field	10^{11}
Urban noise	10^{9}
Abdominal currents	10^{5}
Cardiac activity, skeletal muscle	10^{5}
Cortical evoked fields	10^{2}
SQUID noise	10

Table 3.
Behavioral Data for the MEG Experiment

Condition	Mean RT	SD	SE
Irregular			
Past	808	99.9	16.6
Stem	577	80.4	13.4
Regular			
Past	781	105.2	17.5
Stem	560	70.1	11.7

Note: RT=Reaction Time, in ms

Table 4
RMS Data of the MEG Experiment

Condition	Peak1			Peak2			Peak3			Peak4		
	Mean	SD	SE	Mean	SD	SE	Mean	SD	SE	Mean	SD	SE
	Latency[a]											
Irregular												
Past	231	16.1	2.9	295	30.7	5.5	361	15.8	2.852	431	17.3	3.1
Stem	218	17.3	3.1	269	14.2	2.5	329	16.5	2.8			
Regular												
Past	221	16.9	3.1	293	14.9	2.6	344	20.3	3.6	412	15.1	2.7
Stem	216	16.2	2.9	262	17.6	3.1	326	17.6	3.1			
	Amplitude[b]											
Irregular												
Past	7.8	3.1	0.5	6.7	4.9	0.8	5.8	4.6	0.8	5.2	2.3	0.4
Stem	6.2	3.1	0.6	5.5	1.9	0.3	5.4	3.1	0.5			
Regular												
Past	6.3	2.9	0.5	5.3	2.3	0.4	5.1	2.2	0.3	5.1	2.2	0.4
Stem	7.3	3.2	0.5	6.2	3.5	0.6	5.1	2.9	0.5			

Note.
[a] units are ms
[b] units are 10^{-13} T

Table 5
Statistics of the RMS Data

1. 200-300ms After Onset of Stimulus

Source	df	F-Value
Amplitude		
regularity	1,30	0.697
task	1,30	0.181
time	1,30	9.224*
regularity * task	1,30	13.406*
regularity * time	1,30	0.053
task * time	1,30	0.061
regularity * task * time	1,30	0.183
Latency		
regularity	1,30	3.462
task	1,30	47.675**
time	1,30	1726.144**
regularity * task	1,30	0.073
regularity * time	1,30	0.075
task * time	1,30	20.442**
regularity * task * time	1,30	2.045

* p<.01, ** p<.001

Table 5, continued

2. 300-400ms After Onset of Stimulus

Source	df	F-Value
Amplitude		
regularity	1,30	1.931
task	1,30	0.061
regularity * task	1,30	0.348
Latency		
regularity	1,30	12.356**
task	1,30	54.003***
regularity * task	1,30	4.642*

* p<.05, ** p<.01, *** P<.001

3. 400-500ms After Onset of Stimulus

Source	df	F-Value
Amplitude		
regularity	1,30	0.121
Latency		
regularity	1,30	26.96*

* p<.001

Table 6.
ECD data of the MEG Past Tense Production Experiment

condition	time (ms)	EMSE Coordinates (mm)[1]			Talairach Coordinates (mm)[2]		
		x	y	z	x	y	z
Irregular Past	210-230	-40	48	42	-45	-20	7
	290-310	-35	40	35	-40	-18	4
	350-380	-35	42	30	-42	-18	2
	430-450	-42	40	19	-40	-20	-10
Stem	200-230	-40	40	19	-40	-20	-10
	250-280	-41	35	29	-35	-21	2
	320-340	-45	45	35	-42	-24	4
Regular Past	220-240	-38	40	35	-40	-19	4
	290-310	-32	33	30	-33	-9	2
	330-350	35	38	55	-50	6	22
	410-430	30	35	50	-47	4	18
Stem	200-230	-41	48	25	-45	-21	-7
	250-280	-40	50	31	-49	-20	2
	320-340	-46	35	34	-35	-25	4

Note.

[1] 3-D coordinate system for the EMSE coordinates are as follows: the origin (0,0,0) marks the points defined by the intersection between (a) the line between the tragus landmark on each ear, and (b) a line drawn perpendicular to this line from the nasion. The x-y plane is defined by these two lines.

+x = anterior; -x = posterior
+y = left; -y = right
+z = superior; -z = inferior

[2] very approximate (see text)

Table 7
Statistics on Sensor by Sensor Comparison of the MEG Data
Last Two Peaks, Past Tense Conditions Only

Source	df	F-value
sensor	16,480	3,349***
regularity	1,30	0.12
sensorXregularity	16,480	3.005***
time	3,90	0.275
sensorXpeak	48,1440	1.86***
regularityXtime	3,90	.58
sensorXregularityXtime	48,1440	2.385***

*** p<.001

Table 8
Contrast Values for the fMRI Subtractions

Subtraction	Condition	Contrast Value
All-Fixation	RegularPast	1
	RegularStem	1
	IrregularPast	1
	IrreularStem	1
Past-Stem	RegularPast	1
	RegularStem	-1
	IrregularPast	1
	IrreularStem	-1
RegularPast– RegularStem	RegularPast	1
	RegularStem	-1
	IrregularPast	0
	IrreularStem	0
IrregularPast– IrregularStem	RegularPast	0
	RegularStem	0
	IrregularPast	1
	IrregularStem	-1
(RegularPast- RegularStem) - (IrregularPast- IrregulsrStem)	RegularPast	1
	RegularStem	-1
	IrregularPast	-1
	Irregularstem	1
(IrregularPast- IrregularStem (RegularPast- RegularStem)	RegularPast	-1
	RegularStem	1
	IrreguarPast	1
	IrregularStem	-1

Table 9
RegularPast-Regular Stem, Significant Clusters

x,y,z (mm)	Location	T	(Z)	$P_{uncorrected}$
-48, 6,18	left BA 44/45	5.63	4.12	0.000
27, 36, -6	right BA 47	5.51	4.07	0.000
27, -12, 45	right BA 6	5.45	4.04	0.000
-21, 15, -15	left paraphippocampal gyrus	4.84	3.75	0.000
6, 3, 57	right BA 6	4.77	3.71	0.000
12, 6, 51	right BA 6	4.27	3.44	0.000
12, -6, 30	right BA 24	4.43	3.53	0.000
-27, -3, 63	left BA 6	4.3	3.46	0.000
-21, -57, 42	left BA 7	4.09	3.33	0.000

Table 10
IrregularPast-IrregularStem, Significant Clusters

x,y,z (mm)	Location	T	(Z)	$P_{uncorrected}$
-6, 12, 18	left corpus callosum	5.31	3.98	0.000
-39, 30, -3	left BA 47	4.95	3.8	0.000
-6, 3, 51	left BA 6	4.8	3.72	0.000
-3, 12, 60	left BA 6	4.34	3.48	0.000
-6, 15, 48	left BA 6	3.96	3.26	0.001
-48, 6, 24	left BA 44/45	4.58	3.61	0.000
-36, 6, 27	left IFG	4.03	3.3	0.000
-39, -24, -15	left paraphippocampal gyrus	4.47	3.55	0.000
-21, -63, 33	left BA 7	4.46	3.54	0.000
-48, 39, 3	left 45/47	4.45	3.54	0.000
-45, -42, 39	left BA 40	4.44	3.53	0.000
9, 9, -9	right BA 25	4.34	3.48	0.000
21, -54, 36	right BA 7	4.15	3.37	0.000

Table 11
(RegularPast-RegularStem)-(IrregularPast-IrregularStem), Significant Clusters

x,y,z (mm)	location	T	(Z)	$P_{uncorrected}$
21,-33, 42	right parietal	4.38	3.5	0.000

Table 12
(IrregularPast-IrregularStem)-(RegularPast-RegularStem), Significant Clusters

x,y,z (mm)	Location	T	(Z)	$P_{uncorrected}$
-24. –75, 21	left parietal	4.5	3.56	0.000
6, 24, 48	right anterior cingulate	4.3	3.45	0.000
-42, 33, -3	left BA 45/47	4.24	3.42	0.000
-3, 39, 3	left anterior cingulate	3.93	3.24	0.001
48, 21, -6	right IFG	3.58	3.02	0.001
27, -57, 42	right parietal	3.49	2.96	0.002
-36, 6, 39	left IFG	3.32	2.86	0.002

Table 13
Summary of Hypotheses and Results

Condition	Hypotheses		Data	
	Words and Rules	Connectionist	fMRI	MEG
Past Irregular	Activation in temporal areas (lexical look-up of stem and stored past tense form)	No differences	- Right anterior cingulate - Left frontal clusters entered around [-42,33,-3] and around [-36,6,39] - Right parietal cluster	Temporal areas (entire time course)
Regular	Activation in temporal areas (lexical lookup of stem) **and** Broca's area (application of the rule)		- Right parietal cluster	Posterior areas (early) and anterior areas (late)
Stem	Activation in temporal areas	No differences	N/A	Temporal areas

Bibliography

Awh, E., Jonides, J., Smith, E.E., Schumacher, E.H., Koeppe, R.A., & Katz, S. (1996). Dissociation of storage and rehearsal in erbal working memory. *Psychological Science, 7,* 25-31.

Bergida, R., O'Craven, K., Savoy, R. L., & Ullman, M. T. (1998). fMRI double dissociations in frontal and temporal regions between regular and irregular past tense production. Paper presented at the 5th Annual Meeting of the Cognitive Neuroscience Society, San Francisco, CA.

Binder, J.R., Springer, J.A., Bellgowan, P.S.F., Swanson, S.J., Frost, J.A., & Hammeke, T.A. (1997). A comparison of brain activation patterns produced ay auditory and visual lexical-semantic language tasks. *NeuroImage, 5,* S558.

Broca, P. (1861). Remarques sur le siège de la faculté de la parole articulée; suivies d'une observation d'aphémie (perte de parole). *Bulletin de la Société d'Anatomie (Paris), 36,* 330-357.

Broca, P. (1865). Sur le siège de la faculté du langage articulé. *Bulletin d'Anthropologie, 6,* 377-393.

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

Bullinaria, J.A., & Chater, N. (1995). Connectionist modelling: Implications for cognitive neuropsychology. *Language and Cognitive Processes, 10*: 227-264.

Bullinaria, J.A., & Chater, N. (1996). Double dissociation, modularity, and distributed organization. *Behavioral and Brain Sciences, 19/4*: 632-.

Bunge, S.A., Klingberg, T., Jacobsen, r.B., Gabrieli, J.D.E. (2000). A resource model of the neural basis of executive working memory. *PNAS, 97/7,* 3573-3578.

Caplan, D. (1987). *Neurolinguistics and linguistic aphasiology. An introduction.* Cambridge: Cambridge University Press.

Caplan, D., Hildebrandt, N., & Markis, N. (1996). Location of lesions in sroke patients with deficits in syntactic processing in sentence comprehension. *Brain, 119,* 933-949.

Caramazza, A. (1988). Some aspects of language processing revealed through the analysis of acquired aphasia: The lexical system. *Annual Review of Neuroscience, 11,* 395-421.

Caramazza, A. (1996). Pictures, words and the brain. *Nature, 383,* 216-217.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology, 14/1,* 177-208.

Caramazza, A. (1998). The interpretation of semantic category-specific deficits: What do they reveal about the organization of conceptual knowledge in the brain? Introduction. *Neurocase, 4/4-5,* 265-272.

Caramazza, A., & Hillis, A.E. (1990). Where do semantic errors come from? *Cortex, 26,* 95-122.

Caramazza, A., & Hillis, A.E. (1991). Lexical organization of nouns and verbs in the brain. *Nature, 359,* 788-789.

Caramazza, A., & Miozzo, M. (1997). More is not always better: a response to Roelofs, Meyer, and Levelt. *Cognition, 64,* 309-343.

Chee, M.W., Buckner, R.L., O'Craven, K.M., Bergida, R., Rosen, B.R., & Savoy, R.L. (1997). Auditory and visual word processing studied with fMRI. *NeuroImage, 5,* S557.

Chomsky, N. (1981). *Rules and Representations.* Cambridge, MA: MIT Press.

Chomsky, N. (1986). *Knowledge of Language: the Managua lectures.* Cambridge, MA: MIT Press.

Clahsen, H., Eisenbeiss, S., & Sonnenstuhl-Henning, I. (1997). Morphological structure and the processing of inflected words. *Theoretical Linguist, 23/3,* 201-250.

Clahsen, H. (1999). Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences, 22/6,* 991-+.

Cobuild Corpus, http://titania.cobuild.collins.co.uk/boe_info.html

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M.A., & Michel, F. (2000).The visual word form area - Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain, 123,* 291-307.

Cuffin, N.B. (1985). A comparison of movng dipole inverse solutions using EEG's and MEG's. *IEEE, BME-32/11,* 905-910.

Dale, A.M., & Sereno, M.I. (1993). Improved localization of cortical activity by combining EEG and MEG with MRI ocrtical surface reconstruction: A Linear approach. *Jounal of Cognitive Neuroscience, 5/2,* 162-176.

Damasio, A.R. , & Tranel, D. (1993). Nouns and verbs are retrieved with differently distributed neural systems. *Proceedings of the National Academy of Science USA, 90,* 4957-4960.

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

Demb, J.B., Desmond, J.E., Wagner, A.D., Vaidya, C.J., Glover,G.H., & Gabrieli, J.D.E. (1995). Semantic encoding and retrieval ine left inferior prefrontal cortex: A functional mri sudy of task difficulty and process specificity. *J. Neurosci., 15,* 5870-5878.

Diamond, A., & Goldman-Rakic, P.S. (1989). Comparison of human infants and rhesus-monkeys on Piaget ABBAR tas – evidence for dependence on dorsolateral prefrontal cortex. *Experimental Brain Research, 74/1,* 24-40.

Dronkers, N.F. (1997). A new brain region for coordinating speech articulation. *Nature, 384,* 159-161.

Embick, D., Marantz, A., Miyashita, Y., O'Neil, W., & Sakai, K.L. (2000). A syntactic specialization for Broca's area. *PNAS, 97/11,* 6150-6154.

Fiez, J.A., Raichle, M.E., Miezin, F.M., & Petersen, S.E. (1995). PET studies of auditory and phonological processing: Effects of stimulus characterisics and task demands. *J. Cognitive Neuroscience, 7/3,* 357-375.

Fiez, J.A. (1997). Phonology, semantics, and the role of the left inferior prefrontal Cortex. *Human Brain Mapping, 5/2,* 79-83.

Fiez, J.A., & Petersen, S.E. (1998). Neuroimaging studies of word reading. *PNAS, 95/3,* 914-921.

Francis, N. & Kucera, H. (1982). *Frequency analysis of english usage: lexicon and grammar*. Boston: Houghtn Mifflin.

Friederici, A.D., Opitz, B., & von Cramon, D.Y. (2000). Segregating semantic and syntactic aspects of processing in the human brain: an fMRI investigation of different word types. *Cerebral Cortex, 10,* 698-705.

Gabrieli, J.D.E., Poldrack, R.A., & Desmond, J.E. (1998). The rle of left pefrontal cortex in language and in memory. *PNAS, 95/6,* 908-913.

Garrett, M. (1992). Disorders of lexical selection. *Cognition 42,* 143-180.

Gopnik, M. (1990). Genetic basis of grammar deficit. *Nature, 344,* 715-715.

Gopnik, M., & Crago, M. (1991). Familial aggregation of a developmental language disorder. *Cognition, 39,* 1-50.

Grabowski, T.J., Damasio, H., & Damasio, A.R. (1997a). Premotor and prefrontal correlates of category-specific lexical retrieval. *Neurology, 48/3,* 3046-3046, Suppl. 2.

Grabowski, T.J., Damasio, H., & Damasio, A.R. (1997). Premotor and prefrontal activation during category-specific lexical retrieval. *NeuroImage. 5,* S549.

Gutbrod, K., Nirkko, A.C., Ozdoba, C., Schroth, G., & Schnider, A. (1997). Selective activation of Broca's and Wenicke's area in functional magnetic resonance imaging (fMRI). *NeuroImage. 5,* S549.

Hagoort, P., Indefrey, P., Brown, C., Herzog, H., Steinmetz, H., & Seitz, R.J. (1999). The neural circuitry involved in the reading of German words and pseudowords: A PET study. *Journal of Cognitive Neuroscience, 11/4,* 383-398.

Hahn, U., & Chater, N. (1998). Similarity and rules: distinct? exhaustive? empirically distinguishable? *Cognition, 65/2-3,* 197-230.

Halle, M., & Marantz, A. (1993). Distributed Morphology and the Pieces of Inflection. In K. Hale & S.J. Keyser (Eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger.* Cambridge, MA: MIT Press.

Hari, R., Salmelin, R., MÑkelÑ, J.P, Salenius, S., & Helle, M. (1997). Magnetoencephalographic cortical rhythms. *International Journal of Psychophysiology, 26,* 51-62.

Hillis, A.E., & Caramazza, A. (1995). Representation of grammatical categories of words in the brain. *Journal of Cognitive Neurosciece, 7/3,* 396-406.

Howard, D., Patterson, K., Wise R., Brown, D., Friston K., Weiller, C., & Frackowiack, R.S.J. (1992). The cortical localization of the lexicons. *Brain, 115,* 1769-1782.

Icardo, C., Décorps, M., & Segebarth, C. (1997). Neural correlates of covert language production: An fMRI study. *NeuroImage, 5,* S565.

Indefrey, P. (1997) PET research on language production. *International Journal of Psychopyshiology, 25/1,* 81-81.

Indefrey, P., Brown, C., Hagoort, Pl, Herzog, H., Sach, M., & Seitz, R.J. (1997). A PET study of cerebral activation patterns induced by verb inflection. *NeuroImage. 5,* S548.

Indefrey, P. (1999). Some problems with the lexical status of nondefault inflection. *Behavioral and Brain Sciences, 22/6/,* 1025-+.

Jaeger J.J., van Valin R.D., & Lockwood A.H. (1998). Evaluating behavioral and neuroimaging data on past tense processing - Response to Seidenberg and hoeffner. *Language, 74/1,* 123-128.

Jaeger, J., Lockwood, A.H., Kemmerer D.L., van Valin, R.D., Murphy B.W., & Khalak H. G. (1996). A positron emission tomographic study of regular and irregular verb morphology in English. *Language, 72/3,* 415-497.

Jonides (1990). Memory. In D. Osherson, H. Lasnik, D. Scarborough, & S. Sternberg (Eds.), *An Invitation to cognitive science.* Cambridge, MA: MIT Press.

Jonides, J., Schumacher, E.H., Smith, E.E., Koeppe, R.A., Awh, E., Reuter-Lorenz, P.A., Marshuetz, C, & Willis, C.R. (1998). The role of parietal Cortex in verba working memory. *J Neuroscience, 18/13,* 5026-5034.

Kirchhoff, B.A., Wagner, A.D., Marl, A., & Stern, C.E. (2000). Prefrontal-temporal circuitry for episodic encoding and subsequent memory. *J. Neuroscience, 20/16*, 6173-6180.

Laudanna, A. , & Cermele A., & Caramazza, A. (1997). Morpho-lexical representations in naming *Language and Logical Processes, 12/1*, 49-66.

Lauter, J.L. & Hirsch, I.J. (1985). Speech as a temporal pattern – a psychoacoustical profile. *Speech Communication, 4/1-3*, 41-54.

Levelt, W., Roelofs, A. , & Mayer, A.S. (1998). A theory of lexical access in speech production. *Behavioral Brain Sciences, 22/2*, 1-+.

Liddle, P.F., Kiehl, K.A., & Smith, A.M. (2001). Event-related fMRI study of response inhibition. *Human Brain Mapping, 12/2*, 100-109.

Liu, A.K., Belliveau, J.W., & Dale, A.M. (1998). Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations. *Proceedings of the national Academy of Science USA, 95*, 8945-8950.

MacDonald III, A.W., Cohen, J.D., Stenge,r V.A., & Carter, C.S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science, 288*, 1835-1838.

Marchman, V.A., Plunkett, K., Goodman, J. (1997). Overregularization in English plural and past tense inflectional morphology: a response to Marcus (1995*). Journal of Child Language, 24/3, 767-779.*

Marcus, G., Pinker, S., Ullman, M., Hollander, M., Rosen T.J., & Xu, F. (1992). Overregularization in language acquisition. Monographs of the Society for Research in Child Development. Serial No 228, Vol. 57, No. 4

Marcus, G.F., Brinkman, U., Clahsen, H., Wiese, R., Woest, A., & Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology, 29/3,* 189-256.

Marslen-Wilson, W., & Tyler, L.K. (1997). Dissociating types of mental computation. *Nature 387,* 592-594.

Marslen-Wilson, W., & Tyler, L.K. (1998). Rules, representations, and the English past tense. *Trends in Cognitive Science, 2/11,* 428-435./

Martin, A., Haxby, J.V., Lalonde, F.M., Wiggs, Ch.L., & Ungerleider, L.G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science, 270,* 102-105.

Martin, A., Wiggs, Ch., Ungerleider, L.G., & Haxby, J.V. (1996). Neural correlates of category-soecific knowledge. *Nature, 379,* 649-652.

Mummery C.J., Scott, S.K., Ashburner, J., & Wise, R.J.S. (1997). Both cerebral hemispheres recognise words. *NeuroImage, 5,* S563.

Mummery C.J., Ashburner, J., Scott, S.K., & Wise, R.J.S. (1999) Functional neuroimaging of speech perception in six normal and two aphasic subjects. *Journal of the Acoustical Society of America, 106/1,* 449-457.

Nishitani, N., & Hari, R. (2000). Temporal dynamics of cortical representation for action. *PNAS, 97/2,* 913-918.

Nobre, A., McCarthy, G., & Allison, T. (1994). Word recognition in the human inferior temporal lobe. *Nature 372:,* 260-263.

Nunez, P. (1981). Electric fields of the brain: The neurophysiology of EEG. Oxford: Oxford University Press.

Ochsner, K.N., Kosslyn, S.M., Cosgrove, G.R., Cassem, E.H., Price, B.H., Nierenberg, A,A,, & Rauch, S.L. (2001). Deficits in visual cognition and attention following bilateral anterior cingulotomy *Neuropsychologia, 39/3,* 219-230.

Orrison, W. (1988a). EEG and clinical encephalography. In Orrison, W. jr (Ed.) *Functional Brain Imaging*

Orrison, W. (1988b). Magnetoencephalography. In Orrison, W. jr (Ed.) *Functional Brain Imaging*

Orsolini, M., & Marslen-Wilson, W. (1997). Universals in morphological representation: Evidence from Italian. *Language and Cognitive Processes, 12/1,* 1-47.

Paulescu, E., Frith, C.D., & Frackowiak, R.S.J. (1993). The neural correlates of the verbal component of working memory. *Nature, 362,* 324-345.

Paulesu, E., Signorini, M., Frith, C.D., Menoncello, L., Cappa, S.F., Perani, D., & Fazio, F. (1997). Functional heterogeneity of left inferior frontal cortex: Explicit behavioral and PET evidence. *NeuroImage. 5,* S549.

Penke, M., Weyerts, H., Gross, M., Zander, E., Münte, Th., & Clahsen, H. (1997). How the brain processes complex words: an event-related potential study of German verb inflections. *Cognitive Brain Research, 6/1,* 37-52.

Petersen, S.E., Fox, P.T., Posner, M.I., Mintun, M., & Raichle, M.E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature, 331,* 585-589.

Petersen, S.E., Fox, P.T., Snyder, A.Z., & Raichle, M.E. (1990). Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. *Science, 249,* 1041-1044.

Pinker, S. (1991a). Rules of language. *Science, 253,* 530-535.

Pinker, S. (1991b). Why the Child Holded the Rabbit. In D. Osherson & H. Lasnik (Eds.), *Invitation to Cognitive Science, vol 2: Language*. Cambridge, MA: MIT Press

Pinker, S. , & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of langguage acquisition. *Cognition, 28,* 73-193.

Pinker, S. , & Prince, A. (1994). Regular and Irregular Morphology and Psychological Status of Rules of Grammar. In S.D. Lima, S.D., R.L. Corrigan, & G.K. Iverson (eds), *The Reality of Linguistic Rules.* Amsterdam/Philadelphia: John Benjamins Publishing Company

Pinker, S. (1999). Words and rules: the ingredients of language. New York, NY: Basic Books.

Pinker, S. (2001) Four decades of rules and associations, or whatever happened to the past tense debate? In E. Dupoux (Ed.), Language, the brain, and cognitive development: Papers in honor of Jacques Mehler. Cambridge, MA:MIT Press.

Plunkett, K. , & Nakisa, R.C. (1997). A connectionist model of the arabic plural system. *Language and Cognitive Processes, 12/5-6,* 807-836.

Poldrack, R.A., Wagner, A.D., Prull, M.W., Desmond, J.E.,Glover, G.H., & Gabrieli, J.D.E. (1998). Functional specialization for semantic and phonological processing in the Left Inferior Prefrontal Cortex. *NeuroImage, 10,* 15-35.

Prasada, S., Pinker, S., & Snyder, W. (1990). Some evidence that irregular forms are retrieved from memory but regular forms are rule-generated. Poster presented at the 31st annual meeting of the Psychonomic Society, Nov. 16-18, New Orleans, 1990.

Prasada, S. , & Pinker, S. (1993). Generation of regular and irregular morphological patterns. *Language and Cognitive Processes, 8/1,* 1-56.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral Brain Sciences, 22/2* , 253-+.

Redington, M., & Chater, N. (1998). Connectionist and statistical approacges to language acquisition: a distributional perspective. *Language and Cognitive Processes, 13/2-3,* 129-191.

Rowe, J.B., Toni, I., Josephs, O., Frackowiak, .S.J., & Passingham, R.E. (2000). The prefrontal cortex: Respose selection or maintenance within working memory? *Science, 288,* 1656-1660.

Rubia, K., Russell, T., Overmeyer, S., Brammer, M.J., Bullmore, E.T., Sharma, T., Simmons, A,. Williams, S.C.R., Giampietro, V., Andrew, C.M., Taylor, E. (2001). Mapping motor inhibition: Conjunctive brain activations across different versions of go/no-go and stop tasks. *NeuroImage, 13/2,* 250-261.

Rumelhart, D.E. , & McClelland, J.L. (1986). On learning the past tenses of English verbs. In J.L. McClelland, D.E. Rumelhart, and the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 2: Psychological and biological models.* Cambridge, MA: Bradford Books/MIT Press

Rumelhart, D.E. , & McClelland, J.L. (1987). Learning the past tenses of Enlish verbs: Implicit rules of parallel distributed processing? In B. MacWhinney (Ed.), *Mechanisms of language acquisition.* Hillsdale, NJ: Erlbaum

Salmelin, R., Service, E., Kiesilää, P, Uutela, K., & Salonen, O. (1996). Impaired visual word processing in dyslexia revealed with megnetoencephalography. *Annals of Neurology, 140/2:*, 157-162.

Thayer, J.F., and Lane, R.D. (2000). A model of neurovisceral integration in emotion regulation and dysregulation. *Journal of Affective Disorders, 61/3,* 201-216.

Seidenberg, M.S., & Hoeffner, J.H. (1998). Evaluating behavioral and neuroimaging data on past tense processing. *Language, 74/1*, 104-122.

Stowe, A., Broere, C.A.J., Paans, A.M.J., Wijers, A.A., Koster, J., & Vaalburg, W. (1997). Working memory components in a language task. *NeuroImage. 5*, S552.

Stromswold, K., Caplan, D., Alpert, N., & Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomography. *Brain and Language, 52/3*, 452-473.

Thompson-Schill, S.L., D'Esposito, M., Aguirre, G.K., & Farah, M.J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledg: A reevalutation. *PNAS, 94*, 14792-14797.

Thompson-Schill, S.L., Swick, D., Farah, M.J., D'Esposito, M., Kan, I.P., & Knight, R.T. (1998). Verb generation in patients with focal frontal lesions: A neuropsychological test of neuroimaging findings. *PNAS, 95/26*, 15855-15860.

Thompson-Schill, S.H., D'eposito, M., & Kan, I.P. (1999). Effects of repetition and competition on activity in left prefrontal cortex during word generaton. *Neuron, 23*, 513-522.

Ullman, M., Corkin, S., Coppola, M., Hickok, G, Growdon, J., Koroshetz, W., & Pinker, S. (1997a). A neural dissociation within language: Lexicon a part of declarative memory, grammar processed by procedural system. *Journal of Cognitive Neuroscience, 9/2:*

Ullman, M., Bergida, R., & O'Craven, K.M. (1997b). Distinct fMRI activation patters for regular and irregular past tense. *NeuroImage. 5,* S549.

Ullman, M. (1998). A role for declarative and procedural memory in language. *Brain and Cognition, 37/1,* 142-143.

Vandenberghe, R., Price, C., Wise, R., Josephus, O., & Frackowiak, R.S.J. (1996). Functional anatomy of a common semantic system for words and pictures *Nature, 383,* 254-256.

von Cramon, D.Y., Friederici, A.D., & Kotz, S.A.(1997). Syntax processing: The involvement of mediobasal temporal structures. *NeuroImage, 5,* S547.

Wagner, A.D., Schacter, D.L., Rotte, M., Koutstaal, W., Aaril, A., Dale, A.M., Rosen, B.R., & Buckner, R.L. (1988). Building memories: ememberig and Forgetting of verbal experiences as predicted by brain activity. *Science, 281,* 1188-1191.

Wagner, A.D. (in press). Cognitive control and episodic memory: Conributions from prefrontal cortex. L. R. Squire & D. L. Schacter (Eds.). Neuropsychology of Memory (3rd ed.). New York: Guilford Press.

Wagner, A.D., Paré-Blagoev, Clark, J., & Poldrack, R. (submitted). Recovering meaning: Left prefrontal cortex guides controlled semanic retrieval.

Wernicke, C. (1874). The apahasic symptom complex: a psychological study on a neurological basis. Kohn and Weigert, Breslay. Reprinted in R.S. Cohen and M.W. Wartofsky, (eds.), *Boston Studies in the Philosophy of Science, vol.*

4. Reidel, Boston, Mass.

Appendix I

Complete List of Experimental Items for the Past Tense Production Study

The items used in the Past Tense Production Study are listed below, in their past tense forms. Word pairs 1 through 28 and paris 33-35 are matched for both frequency (of both stem and past tense forms) and pronounceability of the past tense forms, the rest is matched for frequency only.

Match #	Frequency Match Rating	Phonol. Match Rating	Irregular Past	Regular Past
1	1	1	wept	whipped
2	1	1	dealt	sailed
3	1	1	lent	fanned
4	1	1	crept	stripped
5	1	1	bound	drowned
6	1	1	sought	stayed
7	1	1	strode	owed
8	1	1	ground	frowned
9	1	1	sold	rolled
10	1.5	1.5	dug	sprayed
11	1	1	ate	weighed
12	1	1	slept	slipped
13	1	1	bred	vied
14	1	1	taught	tied
15	1	1.5	swam	swayed
16	1.5	1	spent	planned
17	1	1	bent	strained
18	1	1	lost	passed
19	1	1	slid	sighed
20	1	1	swept	stepped
21	1	1	held	pulled
22	1	1	kept	stopped
23	1	1	told	called
24	1.5	1	meant	joined
25	1.5	1	built	failed
26	1	1.5	froze	viewed
27	1	1.5	brought	tried
28	1	1.5	caught	played

Match #	Frequency Match Rating	Phonol. Match Rating	Irregular Past	Regular Past
29	1	2	strung	dyed
30	1	2	slung	glued
31	1	2	sang	shared
32	1	2	rang	poured
33	1	1.5	sent	raised
34	1	1.5	sank	signed
35	1	1.5	stuck	cried
36	1	2	flung	dried
37	1	2	stole	scored
38	1	2	hid	stirred
39	1	2	fed	dared
40	1.5	2	stung	spied
41	1	2.5	felt	seemed
42	1	2.5	spun	roared
43	1	3	swung	stared
44	1.5	3	wrung	stored
45	1	3.5	swroe	cared
46	1	4	clung	begged
47	1	4	threw	talked
48	1	4	fought	changed
49	1	4	flew	jumped
50	1	4	grew	watched
51	1	4	struck	hoped
52	1	4	bought	caused
53	1	4	wrote	used
54	1	4	broke	dropped
55	1	4	ran	walked
56	1	4	bore	snapped
57	1	4	won	helped
58	1	4	shot	urged
59	1	4	spoke	worked
60	1	4	thought	looked
61	1	4	rode	wished
62	1	4	bled	scrawled
63	1	4	took	asked
64	1	4	drove	served

Appendix II
Channels Used for RMS analysis

Time Window	System	Channels
200-300ms	64 channels	3, 4 , 5, 7, 18, 19, 20, 22, 24, 32, 46, 55, 60, 63
	93 channels	3, 9, 19, 23, 26, 28, 32, 36, 41, 42, 44, 62, 65, 71, 79, 81, 83, 84, 91
300-500ms	64 channels	3, 7, 18, 19, 22, 23, 26, 28, 32, 41, 42, 44, 46, 54, 58, 60, 61, 62
	93 channels	3, 7, 9, 18, 19, 41, 42, 44, 54, 60, 61, 62, 65, 66, 71, 79, 80, 81, 83, 84, 87, 91

Appendix III
Verbal Instructions for the MEG Experiment

Thank you for participating!

The experiment has two parts; the following are the instructions for the first half only. The second half is very similar to the first and you will receive a separate set of instructions on the screen inside the MEG machine.

Please read the following carefully.

When you are inside the machine and the doors are closed, we will first take a measurement of the marker coils on your head. This will take about 2 minutes. Then, the screen above your head will display a sign that says: We are about to begin.

Please continue to look at the screen.

First an asterisk will appear in the middle of the screen:

*

Then it will disappear, and a verb will appear in its place:

pack

(Please produce the past tense form of the verb/Please read the verb) as soon as you have recognized it:

packed.

Please speak loudly and clearly. Then the verb will disappear, and a row of asterisks will appear:

****

They will disappear as soon as you start speaking.

(Please pronounce the verb in the past tense form that seems most natural to you, the one you would use in your ordinary speech.)

Try to be as fast and as accurate as possible.

You will first get 10 practice trials, so you can get used to the set-up. You will be warned before the experiment begins in earnest.

Please speak clearly and loudly, and move as little as possible. There will be 5 breaks, during which you can relax a bit. These breaks last for a few seconds only, however, so do not get too relaxed, or move by a large amount.

If you have any questions, please ask them now.

Appendix IV
Verbal Instructions for the fMRI Experiment

Thank you for participating!

The experiment has one part, and the following are the instructions. Please read them carefully.

Once you have been put inside the scanner, and the anatomical scans are over, the following sentence will be displayed on the screen:

Get ready.

Please continue to look at the screen.

First an asterisk will appear in the middle of the screen:

+

Then it will disappear, and a CUE word will appear in its place. There are three CUE words, 'READ', 'PAST' and 'FIXATE'. If the CUE word is 'READ', the task is to read the following word silently. If the CUE word is 'PAST', the task is to produce the past tense form of the following verb silently. If the CUE word is 'FIXATE', the task is to fixate on the following fixation cross.

The CUE word will disappear, and either a verb, or a fixation cross will appear in its place:

'Pack' or '+'

Please follow the instructions provided by the CUE words silently, and as quickly as possible.

Then the verb will disappear, and another fixation cross will appear. When the screen goes blank, the trial is over.
If you were not able to follow the task until that point, please just go on to the next word.

If the pronounce the verb in the past tense form that seems most natural to you, the one you would use in your ordinary speech.

Try to be as fast and as accurate as possible.

You will first get 10 practice trials, so you can get used to the set-up. You will be warned before the experiment begins in earnest.
The experiment will last for approximately 6 minutes. During this time, we would ask you to move as little as possible, especially your head.

If you have any questions, please ask them now.