THE UNIVERSITY OF CHICAGO

FROM MEANING TO WORDS: AN INVESTIGATION OF PAST TENSE VERB INFLECTION IN ENGLISH COMPARING A FORM TO FORM MAPPING TASK WITH A MEANING TO FORM MAPPING TASK

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE HUMANITIES
DEPARTMENT OF LINGUISTICS
AND
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
DEPARTMENT OF PSYCHOLOGY
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

BY
ARIKA OKRENT

CHICAGO, ILLINOIS
JUNE 2004

Acknowledgements

This project is intended to satisfy the requirements of two academic departments in two separate divisions with work partially completed in a lab in yet a third department and division. The only way to pull off a stunt like this is to choose one's committee wisely. My co-chair, Terry Regier, developed this project with me from the beginning. He is one of the best teachers I have ever seen, and under his energetic tutelage, I discovered capabilities I never knew I had. Howard Nusbaum, my other co-chair, though possibly the busiest person at the University of Chicago, always had time for me, and often generously gave me two hour chunks of it to work out complicated details and soothe my neuroses. David McNeill, a mensch for the ages, has supported me as a professor and a friend since my first day on campus, and stuck with me through all my changing interests. Amy Dahlstrom taught me about the crucial and delicate balance between theory and data in linguistics and set an example for how to do it with cool. Steven Small gave me a home and a family in his lab and taught me about science by asking the tough and important questions.

Many others deserve thanks for their roles in the completion of this work: my professors at Gallaudet University, who introduced me to linguistics and got me hooked, Scott Liddell, Ceil Lucas, and Bob Johnson. My friends for their advice and commiseration: Marie Coppola, Chris Corcoran, Alex and Elaine Francis, Katie Gruber, Mary Rose, Julia Simon, Svetlana Soglasnova, and Elisa Steinberg. Incredibly smart (and fun) people at the Brain Research Imaging Center without whom the imaging project could not have happened: E Chen, Fred Dick, Robert Lyons, Jeremy Skipper, Ana Solodkin, Patrick Wong, and Nick Wymbs. Special mention must be made of the highly skilled bureaucracy navigating powers of Bernadette Brogan, Marjorie Wash, and Vanessa Wright.

My time at the University was marked by loss of three wonderful professors, Karen Landahl, James McCawley, and Kostas Kazasis. I learned an incredible amount from them in the short time I knew them, and I am proud to be a drop in the waves of their influence still rippling around the world.

In my first year here, I received a blow that quite well could have stopped me from continuing: the death of my brother David. His incredible intellect came not from a desire for fame or respect from others, but from a true and deep curiosity about the world. He continues to inspire me to maintain that ideal. The subsequent years have not been easy for my parents or my sister Sarah, and I thank them for their support and especially for their patience. Also noteworthy for her patience is my grandmother, Gizella. She wasn't sure what was taking so long, but she always had faith that I would finish.

Finally, I must acknowledge the loving, unflagging support of my husband Derrick. Without him, not only would I be a poached egg, [1] I'd be an eternal graduate student poached egg.

---

1. See Acknowledgements section of *Higgins, Derrick. 2002. A multimodular approach to model selection in statistical natural language processing. Ph.D. thesis, University of Chicago, Chicago, IL.*

Abstract

In everyday language use, people produce inflected forms of verbs by mapping an intended meaning onto a phonological form. They do not necessarily map a phonological stem form onto a phonological inflected form, though such an operation is possible when called for. This thesis argues that the nature of producing an inflected form is altered by the presentation of a stem form, particularly with regard to the regularity of the inflected form.

A current debate on the nature of morphological representation has been framed in terms of two opposing models of the process of regular and irregular inflection. The dual-route model proposes a rule application process for regular inflection and a separate lexical memory retrieval process for irregular inflection. The single-route model proposes one associative process for both types of inflection. Much of the data employed in the arguments for both models are the results of form to form mapping tasks. If the presentation of a stem form affects regular and irregular verbs differently, the data used may only be informative about the form to form mapping task itself, and not about inflection in general.

In order to investigate whether the production of an inflected form from meaning differs with respect to regularity from the production of an inflected form from another form, I develop a task to elicit specific verbs from meaning and from form that allows for experimental control. This task is used in a series of behavioral experiments using response time as a measurement, as well as in a functional magnetic resonance imaging (fMRI) experiment using hemodynamic response as a measurement.

The results of the experiments are that regular past tense verbs are produced more quickly than irregular verbs when a stem form is presented, but not when no stem form is presented. Additionally, in the brain, greater frontal lobe activation is found for irregular verbs when a stem form is presented, but not when no stem form is presented. The results suggest that exposure to a phonological form of the verb to be inflected alters the nature of the inflection process.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The general topic of this thesis is regular and irregular verb inflection, and how producing an inflected verb may be influenced by exposure to the verb stem. The specific thesis is that exposure to the stem form of a verb affects the accessibility of regular and irregular past tense forms differently. Using two kinds of measurements, reaction time and brain activation, it demonstrates that when a person generates the past tense of a verb, there is a difference between regular and irregular past tense forms only when the stem form of the verb has been presented.

There is a current debate in linguistics and cognitive science about how our ability to inflect words should be represented. The debate has been framed in terms of two general classes of models, the dual-route model and the single-route model. The data most widely dealt with in the debate are the regular and irregular past tenses of English verbs (Daugherty and Seidenberg, 1992; Hare and Elman, 1995; Hare, Elman, and Daugherty, 1995; Joanisse and Seidenberg, 1999; Ling and Marinov, 1993; MacWhinney and Leinbach, 1991; Maratsos, 2000; Marchman, Plunkett, and Goodman, 1997; Marcus, 1995; Marcus et al., 1992; Pinker, 1991; Pinker, 1999; Pinker and Prince, 1988; Pinker and Prince, 1994; Plunkett and Marchman, 1991; Plunkett and Marchman, 1993; Prasada, Pinker, and Snyder, 1990; Prasada and Prince, 1993; Rumelhart and McClelland, 1986; Seidenberg, 1992; Seidenberg and Bruck, 1990; Ullman, 1999) though data from other languages have been dealt with as well, especially from German (Clahsen, 1999; Clahsen and Rothweiler, 1992;

Hahn and Nakisa, 2000; Marcus et al., 1995; Nakisa and Hahn, 1996) but also including Arabic (Plunkett and Nakisa, 1997), Hebrew (Berent, Pinker, and Shimron, 1999), Dutch (Baayen, Dijkstra, and Schreuder, 1997), Italian (Orsolini, Fanari, and Bowles, 1998), and Hungarian (Lukacs, Racsmany, and Pleh, 2001).

The proponents of the dual-route model of inflection propose that irregular inflection, i.e., *sing* goes to *sang* in the past tense, is handled in the lexicon, or an associative memory network, while regular inflection, i.e. *walk* goes to *walked* in the past tense, is handled by a system for rule-governed symbol manipulation.

The proponents of the single-route model of verb inflection propose that both irregular and regular inflection can be handled in a single associative network and that the differences between regular an irregular inflection in behavior, learning, and brain damage can be accounted for by the different distributions of regular and irregular forms in the input, such as the type and token frequency of the forms involved and their phonological similarity.

Different types of evidence have been brought to bear on this issue. Behavioral measures of error rate and reaction time (Marchman, 1997; Prasada, Pinker, and Snyder, 1990; Seidenberg and Bruck, 1990; Stemberger and MacWhinney, 1988), as well as measures of electrical activity and blood flow in the brain (Beretta et al., 2003; Indefrey et al., 1997; Jaeger et al., 1996; Lavric et al., 2001; Newman et al., 1999; Penke et al., 1997; Ullman, Bergida, and O'Craven, 1997; Weyerts et al., 1997) have been collected in experiments comparing regular to irregular inflection in both comprehension and production.

In this thesis, I will focus solely on the production of inflected forms. I will argue that the task predominately employed in the experimental designs used so far has possibly biased the data, and that data obtained with a more ecological task show a different pattern of results. The task in question is the mapping of form to form, and the alternative task is a mapping of meaning to form. In form to form mapping tasks, the subject is presented with one word form as a stimulus and must respond with another word form. For example, the stimulus is the form *sing* and the response is the form *sang*. In a meaning to form mapping task, the form *sing* is never presented. Rather the subject must generate *sang* from the semantics of the concepts [sing] and [past].

There is some empirical evidence that behavior with respect to verb regularity is altered by the presentation of a verb stem. Marchman (1988) and Marcus et al. (1992) report that children are less likely to overregularize verbs (e.g., *holded*) in spontaneous language production than in experimental situations where stem forms are used in elicitation tasks. For example, Marcus et al. find in their analysis of the child language data collected by Kuczaj (1977) and Brown (1973) that all the children had a higher rate of overregularization after a past tense elicitation where a stem form was uttered by a parent. They explicitly suggest that the presentation of a stem form creates a bias against producing the correct irregular form because of priming:

> children are being supplied with the stem itself seconds before they are asked to supply the past form (e.g., This is a girl who knows how to swing. She did the same thing yesterday. She ___). This contrasts with naturalistic settings in which children produce a past form for an irregular in response to a mental representation of the verb's meaning plus the feature for past tense; the phonetic form of the stem need never be activated. Thus, experimental elicitation of irregular past tense forms using the stem as a prompt, like parents' leading questions containing the

> stem, are likely to prime the child's representation of the stem form and possibly suppress the irregular past, leading to an increased likelihood of overregularization (Marcus et al., 1992, page 66).

Elicitation tasks that use form to form mapping may be leading to what looks like greater difficulty with irregular verbs. Not only in the child data, but in all studies that use such tasks. Studies of brain damaged people tend to use elicitation tasks where stem forms are presented (e.g., the studies discussed in Ullman et al., 1997). Patterson et al. (2001) report that semantic dementia patients, who have greater difficulty with irregular than regular verbs, "essentially never make errors on past-tense irregular verbs in their spontaneous speech." The spontaneous speech errors they do make occur, "only in conversational situations that rather resemble the verb generation paradigm, because the questioner just happens to have supplied the present-tense form of the verb" (page 273). The trouble with irregular inflection that children and impaired subjects exhibit may reflect an inability to disengage from the phonological form that has been presented as part of the elicitation stimulus, and not necessarily a problem with irregular inflection per se.

These form to form elicitation tasks have also been used with normal adults in order to look at differences between classes of irregular verbs (Bybee and Slobin, 1982; MacKay, 1976) and also in comparing irregular to regular verbs within the context of the dual-route single-route debate (Beretta et al., 2003; Prasada, Pinker, and Snyder, 1990; Indefrey et al., 1997; Jaeger et al., 1996; Seidenberg and Bruck, 1990; Ullman, Bergida, and O'Craven, 1997). Chapter 2 presents the competing models of inflection and reviews the role that form to form mapping experiments have played for them. Chapter 3 presents a behavioral experiment measuring reaction times of subjects producing past tense form from meaning and from stem forms.

Chapter 4 presents an fMRI (functional magnetic resonance imaging) experiment measuring activity in the brain in meaning to form and form to form tasks. Chapter 5 summarizes the conclusions drawn in this work, and suggests some possible directions for future research.

Chapter 2

Inflection and form to form mapping

2.1 The single and dual route models of inflection

Figure 2.1 is a very simple model of what we do when we produce the past tense form of a verb. We begin with a meaning to express—the verb meaning and the tense meaning—and it is transformed into the correct phonological form. A meaning is mapped to a form.

The form to form mapping task used by both single and dual-route proponents implies an extra processing step, where the form of the stem is accessed and that stem form is then transformed into its past form. Figure 2.2 shows what this model looks like for both single and dual-route models.[1]

First, the verb meaning is transformed into a stem form, and then that stem form is transformed into its past form. In the single route model, all stem and past forms are connected to each other within an associative memory network, and the past tense form produced is activated by a combination of the phonological form of the stem and the statistical properties of the history of memory exposure to words as represented in the relative connection weights in the network. In the dual-route model, only irregular past tense forms are connected to stem forms in a memory network. In producing the past tense, the memory network is activated, and if no

1. The models as diagrammed are not explicitly endorsed by particular researchers. They are rather implied by features of the models as described. This issue will be addressed at the end of this section.

Figure 2.1: Simple model of past tense production

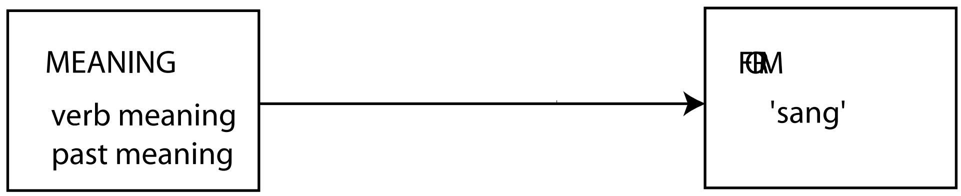
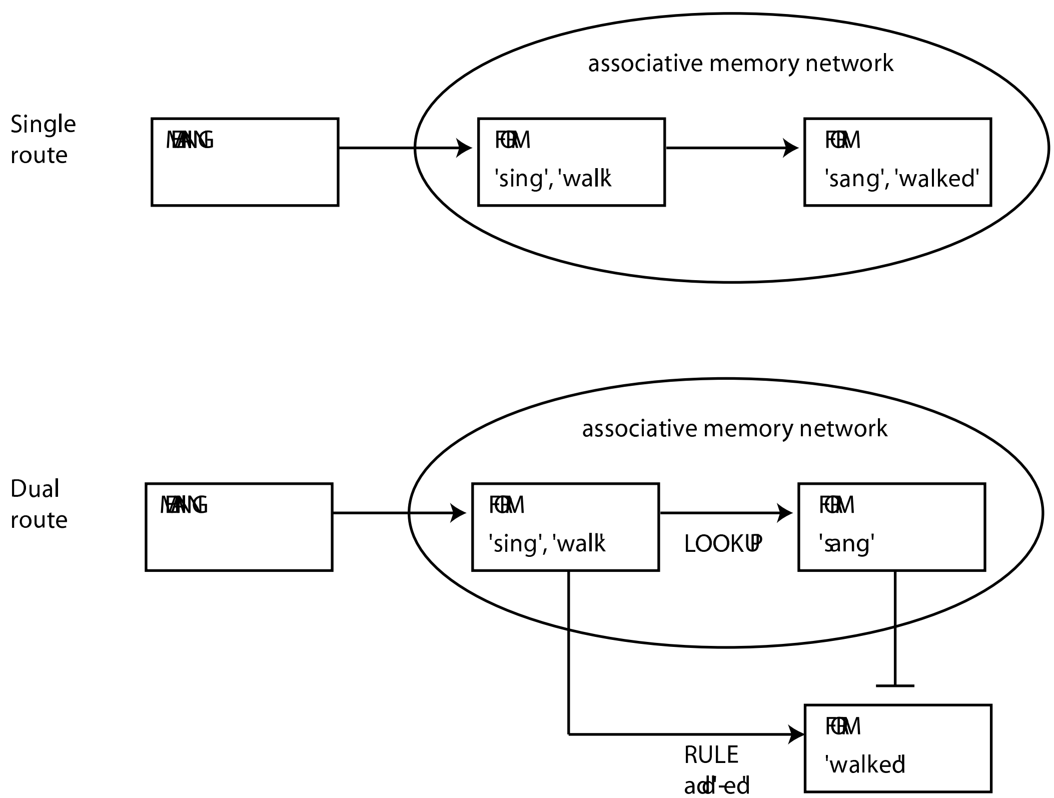


Figure 2.2: Single and dual route models of past tense generation. The circled components illustrate a form to form submapping within the overall meaning to form mapping



past tense form is found in memory, the rule system takes over and produces a regular past tense, If an irregular past *is* found, an inhibitory signal is sent to the rule system.

The two models make different predictions about past tense production. If both regular and irregular past tense are handled in one associative memory system, then differences between them should be based on differences in aspects of their representation in memory, for example, how frequently they are heard. If they are handled in two separate systems, then only irregular inflection should be dependent on memory, and regular processes should be independent of representation in memory. Indeed, one of the main findings of research into this question with the form to form task has been that people take longer to produce low frequency irregular verbs than high frequency ones when presented with the stem, but they take no longer to produce low frequency than high frequency regular verbs. This finding is known as the frequency by regularity interaction. It has been used to support the dual route model, though single route proponents have offered alternative explanations for the finding. The finding will be discussed more thoroughly below.

The models also make different predictions about brain resources used in past tense production. Processes handled by the same cognitive mechanism should be handled in the same brain regions, while processes handled by different cognitive mechanisms should be handled by different brain regions. So if past tense inflection is a dual route process, then regulars and irregulars should activate different brain regions, and if it is a single route process, they should activate the same regions. This simplistic view of the brain/cognition relationship has been challenged (Seidenberg and Hoeffner, 1998, among others), but different patterns of activation for regulars and irregulars have been found using the form to form task, and they have been put forth as support for the dual route model.

Although stem to past mapping tasks have been used to gather evidence to support both models, neither the single nor dual route models have been explicitly formulated as models of production where a phonological stem form must be retrieved and then transformed. Rumelhart and McClelland's (1986) single route model was designed not as a past tense production model, but as a past tense acquisition model, and subsequent models were designed to be able to learn specific facts about past tense not captured in the original model (Hare, Elman, and Daugherty, 1995; MacWhinney and Leinbach, 1991; Plunkett and Marchman, 1993). However, once the training, or acquisition phase of the model is completed, and the networks are tested for the accuracy of their acquired, adult-like representation, they are treated, in essence, as production models. They are evaluated on how well they produce a past tense form from a stem form. There have been a few single route models of inflection that map meaning to form. I will discuss two of them later (Ellis and Schmidt, 1998; Joanisse and Seidenberg, 1999). Cottrell and Plunkett (1995) were mainly concerned with the effects of similarity in input and output structures and not with differences between regulars and irregulars.

The dual route model of Pinker (1991) is not intended to require the retrieval of a phonological stem form before inflection. The regular rule "concatenates an affix with a *variable standing for a stem*" [italics mine] (page 531), not necessarily with a full phonological form. However, phonological properties of the stem must be retrieved in order to implement the search for irregular matches in the memory network, even when the verb is regular. Pinker and Prince (1994), in explaining the Blocking principle, whereby the past tense rule is inhibited when a match in memory is found, give a brief description of how the model behaves in on-line production, "A stem is matched against the memory in which irregulars are stored, and fed into the regular rule mechanism, in parallel (page 344)." And the stem being matched against memory must have a phonological form. Pinker (1999) says explicitly that in a memory search, "The phonemes and syllables in a word contact their counterparts in memory piecemeal, more and more of them finding a match as the milliseconds tick by (page 130)." Therefore, the phonological form of the stem must be retrieved before it is inflected by rule or by lookup.

2.2 Problems with stem to past mapping

Regardless of whether inflection models intend to implicate a stem form to past form mapping in production, the tasks used to support the models are form to form mapping tasks. There are problems with the form to form task and the model of verb production it implies. The first problem with the form to form task is that it is unclear whether it is a linguistic, as opposed to a meta-linguistic, task. Most people know how to explicitly relate forms to each other, and the domain of this activity is generally the classroom. The student of Spanish can transform the infinitive into the third person present; the student of Latin can transform the nominative of a noun into the dative. However, it is far from clear that in spontaneous language production we must call up the stem form of a verb and transform it into the past tense. In everyday speech we may or may not map from forms to other forms, as assumed in the models in Figure 2.2, but we certainly map from meanings to form, as depicted in Figure 2.1. A form to form task may only investigate our metalinguistic ability to transform words into other words. A meaning to form task is more naturalistic and makes fewer a priori assumptions.

The second problem with the form to form mapping task is that the form presented as a stimulus may be either more or less easily transformed into the form expected as a response. To respond with *sang* to *sing* may be easier than to respond with *went* to *go* because the first response is more similar to its stimulus. In the past tense production studies on English, the form presented as stimulus is completely contained within the response form for regulars (*walked* contains *walk*) but not necessarily for irregulars (*sang* does not contain *sing*). The nature of the phonological transformation required is confounded with the variable of interest—regularity. In a meaning to form task, no phonological bias is given with the stimulus.

The meaning to form task I use for the present study is both more naturalistic than the form to form tasks used previously, and also avoids a phonological bias. Subjects are presented with two context establishing clauses, followed by a third clause in which they must fill in the blank with the appropriate past tense verb. For example, to elicit the irregular past tense *drunk*, the following stimulus is presented: *His throat was dry. After making some lemonade, he ___ it*, and to elicit the regular past tense *played*: *They were eager to hear his new composition. So sitting down at the piano, he ____ it.*. In this meaning to form task, the verb itself and the fact that it should be in the past tense are both motivated by the preceding context. No stem form is given, so the subject is not biased toward a particular phonological form. The subject must retrieve an appropriate past tense form from meaning, not transform one form into another, as in the form to form task. Strategies for dealing with problems with this task, such as how to control for variability in how well the contexts predict their target words, will be addressed in the description of the design of the behavioral study in section 3.1.2.

2.3 Review of previous studies

This section first reviews the production studies that used a form to form task. This is followed by a discussion of a few studies that employed a meaning to form task, and an explanation of why further meaning to form study is necessary in order to answer the question of whether form to form tasks create a phonological bias and thus fail to discover facts about our linguistic ability to inflect, rather than facts about our metalinguistic ability to relate forms to each other.

The initial articulation of the single route model of inflection was presented by Rumelhart and McClelland (1986). The task performed by their model was a form to form mapping task. During training, groups of input units representing stem forms in a distributed fashion were associated with output units representing past tense forms. Performance was tested by presenting stem forms to the input units and seeing what the output unit response was. Studies had shown that children acquiring the past tense inflection could generalize the regular $-ed$ inflection to new words (Berko, 1958) and that they first produce irregular verbs correctly and then overregularize them by adding $-ed$ (e.g., *holded*) before finally producing all past tenses correctly (Kuczaj, 1977). The data from child language acquisition was taken as evidence for an innate language acquisition module that learns abstract symbolic rules (Pinker, 1984). Rumelhart and McClelland wanted to show that similar acquisition behavior could be displayed by a connectionist network with no innate language learning module and no representation of abstract rules. To quote them:

> We propose an alternative to explicit inaccessible rules. We suggest that lawful behavior and judgments may be produced by a mechanism in which there is no explicit representation of the rule. Instead, we

> suggest that the mechanisms that process language and make judgments of grammaticality are constructed in such a way that their performance is characterizable by rules, but that the rules themselves are not written in explicit form anywhere in the mechanism (Rumelhart and McClelland, 1986, page 217).

The network was successful in learning both regular and irregular past tense inflection, and it seemed to behave in accordance with the child data. Pinker and Prince (1988) responded with a damaging critique of some important aspects of the model's implementation and training. Among other criticisms, they pointed out that the model's exposure to words in training did not reflect the pattern of exposure that children receive, that the facts about children's overregularization behavior that it captured were wrong, that it was capable of inducing rules not found in any natural language, and that it was incapable of representing many of the normal facts of English inflection, such as different inflections for homophonous words and extension of regular inflection to novel words. Subsequent models were modified in order to address the criticisms (Plunkett and Marchman, 1991; Cottrell and Plunkett, 1995; Daugherty and Seidenberg, 1992; Hare, Elman, and Daugherty, 1995; MacWhinney and Leinbach, 1991), and the debate continues.

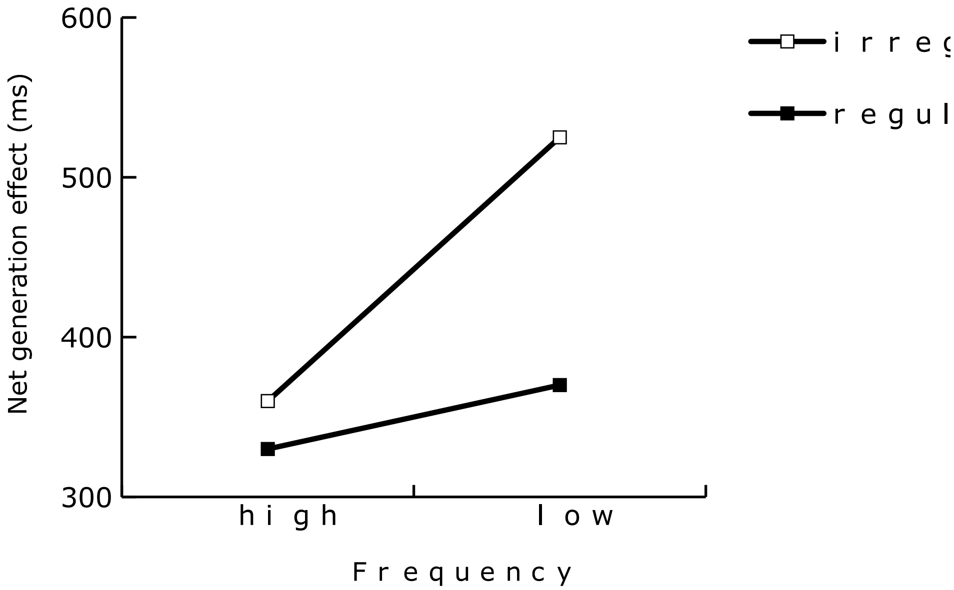
2.3.1 Behavioral studies

Behavioral studies on adults have also been brought to bear on this question. One type of behavioral evidence used to argue for a dual-route model is the relative quickness with which subjects respond with the past tense form when presented with the stem form. In one such experiment Prasada et al. (1990, described in Seidenberg, 1992) presented subjects with a written verb stem like WALK and the subject had to respond as quickly as possible by producing the past tense form *walked*. Both regular and irregular verbs were given. They found that while subjects were faster to produce high frequency irregular past tense forms like *took* than low frequency ones like *bent*, both high and low frequency regular past tenses like *looked* and *basked*, respectively, were produced with the same reaction time (controlling for frequency of stem form). This was interpreted to mean that irregular past tense forms are stored in memory, and since the more often a word is heard, the stronger its memory trace will become, more frequent past tenses will be quicker to retrieve. Because the frequency of a regular past tense form has no effect on retrieval speed, it is assumed that regulars are not stored in memory, but created by rule on-line. Pinker and Prince (1994) offer the Prasada et al. results as support for their theory that for regulars, "prior exposure to and storage of the inflected form is not necessary and affords no crucial advantage. Since memory traces get stronger with additional exposures, the theory predicts that irregular past tense formation should be highly sensitive to frequency, but regular past tense formation in general should not be" (page 327).

Seidenberg and Bruck (1990, described in Seidenberg, 1992) performed a similar experiment and found the same frequency by regularity interaction. There was a significant effect of frequency for irregular verbs, but not for regular verbs (though the low frequency regulars were still slightly slower than the high frequency ones). The effect is shown in Figure 2.3.

However, they do not interpret their results to mean that irregulars are fundamentally different from regulars. A similar frequency by regularity interaction had been observed in naming latencies for regularly and irregularly *spelled* words (MINT has a regular orthography to sound mapping, while PINT is irregular), and

Figure 2.3: Net generation effect (past production latency - stem naming latency) from Seidenberg and Bruck (1990)



Seidenberg and McClelland (1989) had simulated that interaction in a single-route connectionist implementation of orthography to phonology mapping. The frequency by regularity interaction is seen as a consequence of the statistical properties of the input output pairs. Regular verbs benefit not only from exposure to their own input-output pairs, but also from exposure to all the other regular input-output pairs. This is because the weights in the entire network reflect changes made after exposure not only to the input output pairs themselves, but also to many other pairs where $-ed$ occurred in the output units. The connection weights in the network as a whole are modulated to represent the $-ed$ ending as the most likely past tense when a particular input is not strongly associated with a particular output. Therefore, even low frequency regular verbs benefit from the overall high frequency of the regular pattern type. Irregular verbs may benefit from small clusters of other irregulars that make similar input-output mappings (i.e. *sing-sang* benefits from *ring-rang*), but these clusters are not large enough to augment individual frequency effects as much.

The frequency by regularity interaction has been interpreted differently by dual and single route proponents, but both have found that frequency appears not to matter for the production of regular verbs. To dual route proponents, this suggests that regulars are not stored in memory as irregulars are, but rather created on-line by a rule. To single route proponents it suggests that regulars are stored, but the aggregate effects of regulars on the network as a whole give an advantage to low frequency regulars. I will suggest that the finding of the interaction itself may be an artifact of the form to form mapping task, and it is the nature of the task that gives an advantage to low frequency regulars.

The studies described above found an effect of frequency on irregulars, but not on regulars. Two studies using different tasks have found that there are frequency effects for regulars (Alegre and Gordon, 1999; Stemberger and MacWhinney, 1988). They did not look at irregular verbs in the same study, so they did not find an interaction and thus cannot be directly compared with the above studies. What is important is that a frequency effect for regulars has been found, and the studies above did not find such an effect.

Alegre and Gordon (1999) looked at frequency effects not for past tense verbs, but for (regularly) inflected words in general (regular past, 3^{rd} p. present, plural, gerund) in order to test whether inflected forms are stored as whole words. They used a lexical decision task where subjects were presented with an inflected word and had to respond as quickly as possible as to whether it was a word or not. They found that the latency of response to inflected words varied reliably with their inflected word frequencies indicating that inflected words are stored in memory. The production studies above found no frequency effects for regulars, and this may be due to the form to form task demands.

A study by Stemberger and MacWhinney (1988) also found a frequency effect for regulars. They used a task where the subject was given a regular verb in the frame *was* ___*ing* and was asked to speak the past tense of that verb as quickly as possible. All verbs in this study were regular. They did not measure reaction times, but errors. Within regulars, there were significantly more errors on low frequency regulars than on high frequency regulars. This suggested that high frequency regular words may be stored, and not computed by rule on-line. However, Ullman (2001) points out that the regular verbs used in this study were all ‘inconsistent’ regulars—they resemble clusters of irregular verbs (e.g., *glide* is regular, but it rhymes with *slide* and *ride*, which are irregular). In recent versions of the dual-route model (Pinker and Prince, 1994; Ullman, 1999; Pinker, 1999), inconsistent regulars are considered likely to be stored in memory (a position which weakens the theory considerably), so results based on inconsistent regulars are not taken as evidence against the dual-route model.

2.3.2 Imaging studies

The form to form task for past tense verb generation has also been used in brain imaging studies, two in English (Jaeger et al., 1996; Ullman, Bergida, and O’Craven, 1997) and two in German (Beretta et al., 2003; Indefrey et al., 1997). These studies all claim that their results are evidence for a dual-route model of verb inflection because they find non-overlapping areas of cortical activation for regular and irregular past tense verbs. The results from these studies do not agree with each other in which areas are activated by regulars and irregulars, but they all find that irregulars produce more activation than irregulars overall, and that frontal areas in particular are more activated or exclusively activated by irregulars when compared to the activation produced by regulars.

There is an explicit hypothesis by Ullman (1997, 2001) that regular inflection is processed in the left inferior frontal regions and basal ganglia, while irregular inflection is processed in the left temporal region. This hypothesis stemmed from a study comparing the errors made by patients with frontal area damage (anterior aphasia and Parkinson's disease) with those made by patients with temporal area damage (posterior aphasia and Alzheimer's disease) on past tense verb generation. He found that frontal area damage was associated with problems with regular, but not irregular verbs, and temporal damage was associated with problems with irregular, but not regular verbs, a classic double dissociation. It should be pointed out that the brain damage studies used a version of the form to form mapping task. Patients read aloud sentences like *Every day I dig a hole. Just like every day, yesterday I ____ a hole* and were required to fill in the blank. When producing the past tense form, they had very recently read its stem form.

In Ullman's hypothesis, the inferior frontal areas are associated with regular inflection because they are implicated generally in tasks that require a procedure to be executed. This is in contrast to the temporal areas that are implicated in tasks that depend on declarative memory. He does not take the classical modular stance, which would have these areas dedicated to linguistic processes exclusively. Their more general functions are rather suited to particular aspects of linguistic processing. He does, however claim that the handling of regulars and irregulars in functionally distinct areas implies cognitively distinct subsystems: a rule system for regulars and an associative memory system for irregulars.

The imaging studies done so far, including Ullman's own (Ullman, Bergida, and O'Craven, 1997) do not confirm his theory. Jaeger et al. (1996) performed a PET study in which subjects were shown the written form of a verb stem and had to respond with the past tense form. Regular and irregular verbs appeared in different blocks due to the restrictions of PET methodology; single events cannot be analyzed, only blocks of time. The irregular verbs produced more activation overall than the regular verbs. Both regulars and irregulars activated frontal and temporal areas, but an area in the left superior frontal gyrus (BA10) was activated only by irregulars. Citing the fact that this area is implicated in tasks where a response inhibition is required (Cummings, 1993; Goldman-Rakic, 1987; Oscar-Berman, McNamara, and Freedman, 1991), they propose that for irregulars, it functions to suppress the incorrect regularization of the stem. Other results were that irregulars exclusively activated the left primary visual cortex (BA 17) and that regulars exclusively activated the anterior cingulate and the left lateral frontal lobe (BA 46). The method of blocking caused complications in that the regular and irregular tasks differed in ways not addressed by the experiment. For the regular condition, subjects had to do the same thing on every trial—add $-ed$ to the stem, while the irregular condition required a different response on every trial (*sing-sang*, *sleep-slept*, *make-made*). So results may show effects of difficulty or subject expectation rather than of regularity. For a thorough critique of this study see Seidenberg and Hoeffner (1998).

The Indefrey et al. (1997) study of German regular and irregular preterites and participles was also a PET study, also blocked, and also found areas where regulars and irregulars did not overlap. They were not the same areas found by Jaeger et al. (1996). Regulars activated the tempo-parietal areas (right inferior temporal gyrus (BA 37) and left angular gyrus (BA 39)) and irregulars activated frontal areas bilaterally (right inferior (BA 47), left and right middle (BA 44, 6, 46) and left and right superior (BA 9, 10)) in addition to a few occipital and temporal areas. This result is the opposite of Ullman's prediction that regulars are handled frontally and irregulars temporally. This study differed from Jaeger et al. in that it used German, it used a slightly different elicitation task (infinitives were to be inserted in the sentence frame *He* [verbed] *something*), and in analysis regular and irregular activation patterns were compared directly to each other instead of each to a baseline task of reading stems aloud. It is not clear which of these factors contributed to the differing results. The results do agree with the Jaeger et al. study in that more activation was found for irregulars in general, and that the superior frontal gyrus was preferentially activated by irregulars.

Ullman, Bergida, and O'Craven (1997) was an fMRI study that used the same task as Jaeger except that subjects did not speak the response aloud, but covertly to themselves. Analysis compared regular and irregular activation patterns with a baseline task of looking at a fixation point. As in the above studies, the two conditions were blocked. They found that both regulars and irregulars activated inferior frontal regions bilaterally, as well as left temporal areas (they are not more specific about which areas). Frontal activation was greater for irregulars, and temporal activation showed a decreased signal with respect to baseline for irregulars. These results are the exact opposite of the predictions of Ullman's hypothesis. They suggest that the regions might be active in both conditions, but that those regions are playing different roles in the two conditions. This explanation is post hoc and it is unclear how it could be tested. As in the other studies, frontal activation was greater for irregulars. They suggest that for irregulars, frontal activity is related to a lexical search process, but do not give any support for this explanation.

The development of the technique of "single-trial" or "event-related" imaging paradigms, made it possible for experimental conditions to be presented in random order rather than in blocks (Buckner et al., 1996; Josephs, Turner, and Friston, 1997; Zarahn, Aguirre, and D'Esposito, 1997). The study of German regular and irregular nouns and verb participles by Beretta et al. (2003) in using an event-related design, is the only one to avoid the confounds caused by blocking regular and irregular stimuli. Conditions were presented in random order. Subjects were presented with the stem of a noun, or the infinitive of a verb, and instructed to produce the plural of the noun or the participle of the verb covertly. Results from Beretta et al. (2003) that differ from previous studies are that irregulars are more active in the right precentral gyrus, left and right supramarginal/angular/superior parietal lobule and the right posterior temporal lobe. It is like the other studies in that it finds that irregulars produce more activation than regulars overall and that the left prefrontal cortex in particular is more active for irregulars. Their explanation for the greater amount of frontal activation for irregulars is that irregulars place greater demands on working memory. They cite Miller and Cohen (2001), Bunge et al. (2001), and Conway et al. (1999) as evidence that working memory processes rely on the same frontal neural structures as inhibitory processes. In the dual-route model, the production of irregulars should place greater demands on working memory (because output must be monitored and matched against the product of a lexical memory search). If working memory shares neural resources with inhibitory processing, inhibiting a regular response is costly, resulting in greater activation in the prefrontal cortex. Basically, they offer the same explanation as Jaeger et al. (1996) for the greater frontal activation (response inhibition), but in terms that are more amenable to the dual-route model's claim that irregulars involve more memory processing.

The results of the imaging studies performed so far have all been interpreted as evidence for a dual route model because different patterns of activation were found for regular and irregular inflection. It is not clear that separate areas of activation implicate separate rule and memory based processes. The double dissociation in brain damage studies that provides the basis for Ullman (2001) theory has been addressed in connectionist models of regular and irregular spelling-sound correspondences that incorporate semantic units. The double dissociation can result in such models from 'damage' to the connections to phonological or semantic units. Joanisse and Seidenberg (1999), whose connectionist model includes a representation of semantics as well as phonology, found that damage to phonological connections led to greater error rates on regular than irregular verbs, and damage to semantic connections led to the opposite pattern. It may be that a finding of distinct locations for regular and irregular inflection in the brain reflects the relative importance of phonological or semantic processing in the two cases and not two separate production systems, and the findings so far do not determine this.

In all of the tasks discussed so far (except for Alegre and Gordon's, 1999, lexical decision task) the subject has produced the past tense of a word after having been given some form of the stem. In doing this task, the subject is performing a form to form mapping. This task creates a bias in favor of regular verbs in that the form of the stem overlaps completely with that of its past tense forms for regulars (the form *walk* is contained in *walked*) but not for irregulars (the form *sing* only partially overlaps with *sang*). [2] When a subject is given a regular form, a phonological

2. Beretta et al. (2003) made an attempt to deal with this issue by using a class of irregular verbs in German that share as much in common with their stems as regulars do (e.g., the past participle of *laufen* ('to run') is *gelaufen*). However another class of irregular

representation is activated which can be maintained in the production of the past tense. When given an irregular form, the phonological representation activated must be disengaged to a certain extent in order to arrive at the correct phonology for the past tense. In effect, the task of producing a regular past tense is easier because the subject has already had most of its phonological form activated by the cue. Thus the data collected from form to form tasks may not reflect relative differences between regular and irregular verb inflection in linguistic production, but relative differences in task demands. The results may be artifacts of experimental methodology.

2.3.3 Priming

In psycholinguistic terms, the facilitation of a response to a word caused by previous exposure to that word, or to a related word, is called priming. With respect to the lexical decision task most commonly used in studies of inflection, this means that a person more quickly decides that a given stimulus is a word, if they have already seen that word, or a related word. An effective prime facilitates access to the stimulus word.

Previous research on inflection and priming has found that a regular past tense word like *walked* primes its stem *walk* just as well as *walk* primes itself (Fowler, Napps, and Feldman, 1985; Kempley and Morton, 1982; Marslen-Wilson, Hare, and Older, 1993; Napps, 1989; Stanners et al., 1979). The studies disagree on to what degree irregular past tense words prime their stems. Fowler, Napps, and Feldman (1985) found the same degree of priming for regulars and irregulars. Stanners et al.

---

verbs was also included in the study where there is a stem change in the participle (e.g., the past participle of *stehen* ('to stand') is *gestanden*). These two type of irregulars were grouped together for purposes of analysis, so the influence of the phonological change from infinitive to past participle on the results is not ruled out.

(1979) found that irregulars primed their stems, but not as much as regulars did. Kempley and Morton (1982) found no priming at all for irregulars and Marslen-Wilson, Hare, and Older (1993) found an interference effect for a specific class of irregulars.

Though the matter of whether irregular verbs prime their stems has not been settled, there is no question that regular verbs fully prime their stems. Could it be because their stem forms almost fully overlap with their past forms? This possibility is commonly rejected because studies have shown that form overlap is not enough to cause priming in the lexical decision task (e.g., *card* does not prime *car*) (Murrell and Morton, 1974; Kempley and Morton, 1982; Marslen-Wilson et al., 1994). Instead priming is attributed to an abstract morphological level. Gonnerman (1999) however points out that word pairs like *card-car* are far apart in meaning, and shows, in a study of derivational morphology, that form priming does occur when both meaning and form are controlled jointly. Though *trivial* and *trifle* are unrelated morphologically, they overlap in both form *and* meaning, and there is significant priming between them. Priming effects do not arise from relationships in meaning, form, or morphology alone, but from the interaction between meaning and form. Both regular and irregular past tenses overlap with their stems in meaning to the same degree (they are both [stem meaning] + [past]), but they overlap in form with their stems to varying degrees (*walked-walk*, *sang-sing*, *slept-sleep*, etc.). Since the meaning relationship is held constant, form relationships are the likely source of priming effects. Because regulars have a closer form relationship to their stems, they are better primes.

The claim being made here is that the presentation of a regular stem primes its past tense form to a greater degree than an irregular stem primes its past tense form, and that the source of the greater priming effect is the greater phonological overlap between regular stems and their past tenses.

Results of the investigation of the regular-irregular verb question up to this point may have been significantly biased due to the use of a form to form mapping task. Because of the possible phonological bias introduced by the use of a stem form as the stimulus, previous results may reveal facts about how relatively easy or difficult it is for people to disengage from an active phonological state in order to obtain another instead of how people produce regular and irregular verbs. The studies presented in this thesis compare the form to form mapping task with a meaning to form mapping and find that regulars and irregulars differ only in the form to form task.

2.3.4 Meaning to form mapping studies

It should be noted that the proposed study is not the first to use a meaning to form task. There have been two studies of past tense inflection production that employed such a task. One of the studies reported in Stemberger and MacWhinney (1988) was an analysis of naturally occurring error data from spontaneous speech. They counted the number of (no-marking) errors made on verbs that were clearly intended to be produced in the past tense and found a frequency by regularity interaction. A significantly higher proportion of errors were made on low frequency than on high frequency irregular verbs, but no significant difference was found between errors made on low versus high frequency regular verbs, although the effect was in the direction of low frequency having more errors ($p < .10$).

This study of naturally occurring speech errors analyzed, in essence, a meaning to form task. The words were produced in spontaneous conversation and there was not necessarily any pre-given phonological state to influence the results. Yet they found that frequency affected irregular, but not regular verbs in terms of the percentage of incorrect past tenses produced. However the error rate for verbs in general was low (.17%), the result for regulars was in the right direction (more errors on low frequency verbs), and the difference between high and low frequency verbs was marginally significant ($p < .10$). Additionally, only one type of error was analyzed (no-marking), and it is possible that some of the errors were preceded by the production of a stem form in the conversation. So the results of this meaning to form study do not rule out further empirical investigation.

Another meaning to form task was used in an artificial language study by Ellis and Schmidt (1998). They created a mini artificial language of 20 nonsense words to refer to common objects. They trained subjects to name pictures of these objects with the nonsense words. The nonsense words were phonologically similar to the actual words for those objects in order to make them easier to learn (*car* = *garth*, *umbrella* = *brol*). When subjects had learned the vocabulary to a 100% correctness criterion, they were trained on 'plural' forms of the words, by pairing a pair of the same picture with a prefixed form of the stem they had learned previously. Half of the plurals had a regular plural prefix bu- (*bugarth* = *cars*) and half had irregular prefixes (10 different CV syllables for the 10 remaining words). In plural training, frequency of exposure was manipulated so that regular and irregular plurals were precisely crossed with high (five presentations per block) and low (one presentation per block) frequency of exposure. Subjects also reached a 100% accuracy criterion on plural naming. Error and reaction time data showed a frequency by regularity interaction effect later in learning. They attribute these results to the "power law of practice" (Anderson, 1982); there is no difference between high and low frequency regulars at the end of training because the high frequency words have ceilinged out with respect to error rates and floored out with respect to reaction time (the high frequency regulars simply can't get any better). There are effects of regularity in high frequency words, but they are crowded close together by the asymptotic learning function and so are less robust in the face of random error.

In this study the low frequency regulars are not biased toward better performance by a pre-given phonological state because the task required a meaning to form and not a form to form transformation. However, it is unclear how applicable the results of this study are to natural language representation. Artificial language studies have the benefit of offering absolute control over factors of interest that may be confounded in natural language, but what they gain in experimental control, they lose in ecological validity. The toy language of the Ellis and Schmidt study differs from English in some important respects besides the obvious features of size and complexity. Words were learned as responses to pictures, and had no connection with a wide array of conversational contexts over time, as real vocabulary does. Also, plural training stopped after 13–15 training blocks, so it is possible that with further practice, the less frequent irregulars would reach the same asymptote performance level as the other categories, eliminating the frequency by regularity interaction. Additionally, the Ellis and Schmidt results "concern the *learning* of morphology [emphasis theirs]" and not necessarily the processing of morphology (page 315).

In the same study, Ellis and Schmidt presented a single-route computational model of inflection, and confirmed that their human results could be obtained with their single-route model. Where the previous computational models discussed in Section 2.1, learned to convert form input (the stem) into form output (the past tense), this model learned to convert meanings to forms. Representation of both meanings and forms were localist (a unique unit for each meaning, stem, and plural affix), as opposed to distributed (multiple semantic or phonological feature units co-activated for each meaning or word). Because the model did not include a representation of phonology, no phonological effects could be discovered.

Joanisse and Seidenberg (1999) developed a more complex model of inflection in order test whether the double dissociation behavior discovered in studies of brain damaged patients and used to argue for a dual-route account of inflection (Ullman et al., 1997; Ullman, 2001) could be obtained in a single-route model. While meaning in this model was represented in a localist fashion, forms were given a distributed phonological representations. The model was trained to learn four kinds of mapping: meaning to form, form to meaning, stem form to past form, and identity (one form to itself). It was tested on the stem form to past form mapping. When brain damage was simulated by severing phonological connections, performance on the generation of regular verbs deteriorated, and when semantic connections were severed, performance on irregular verbs deteriorated. From these results they argue that it is not necessary to posit separate rule processing and memory processing systems for regular and irregular inflection in order to explain the results from studies of brain damage. Both regular and irregular inflection can depend on a single network of interconnected semantic and phonological representations. The inclusion of a semantic representation and a distributed phonological representation in this model allow the comparison of a meaning to form task with a form to form task in order to see whether regulars and irregulars differ when there is no pre-given phonological state. However, because this study was designed to account for specific behavioral results obtained with a form to form task, no tests were performed in order to see whether regulars and irregulars differed in a meaning to form mapping task.

To sum up, the issue of whether irregular past tense verbs are produced with a qualitatively different mechanism from that used for regular past tense verbs has been investigated using form to form mapping tasks. Behavioral results from these investigations indicate that there is an interaction between frequency and regularity in the time it takes to produce regular and irregular past tenses where frequency affects irregulars, but not regulars. Brain imaging results are conflicting but show that regular and irregular inflection activates some non-overlapping cortical areas. The results found in these studies may not necessarily reflect the mental processes underlying verb inflection, but rather some peculiarities of the form to form task used. The presentation of a stem form like *walk* sets up a phonological state which is completely contained in *walked* while a stem form like *sing* is only partially contained in *sang*. Production reaction times and neurological response may reflect a difference between the type of phonological transformation required, and not a fundamental type difference between regulars and irregulars.

Chapter 3

Behavioral experiments

3.1 Introduction

In this chapter I describe four experiments designed to test the hypothesis that generating the past tense from a stem form is different from generating the past tense directly from meaning. A sentence completion task is used, in which subjects produce specific words from contexts formulated to constrain the selection of those words with high probability. In form conditions, a stem form is presented, while in meaning conditions no stem form is presented. The measurement of interest is response time to produce the target word. The response times to produce regular past tenses are compared with those for production of irregulars, and the difference between regulars and irregulars for form conditions is compared with that for meaning conditions.

3.1.1 Overview of experiments

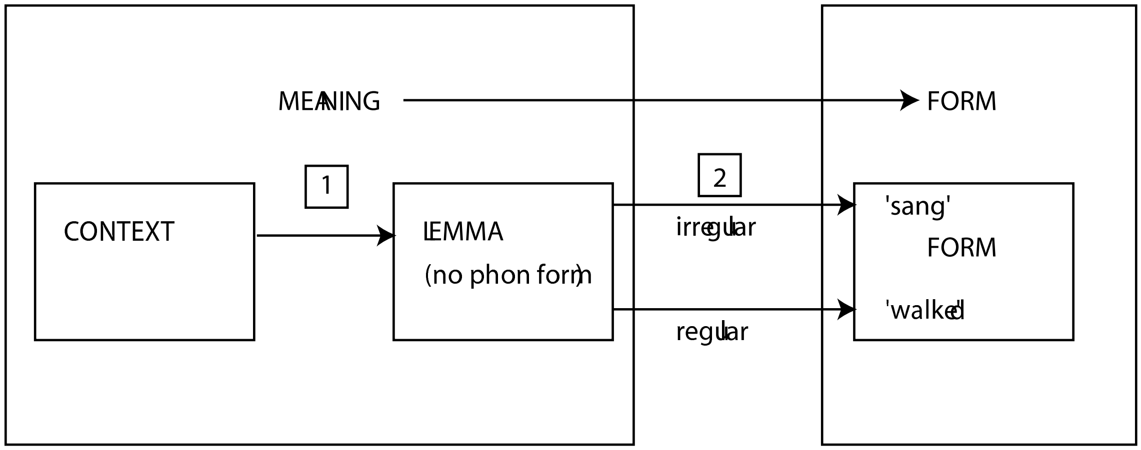
In the first experiment a task in which subjects produce past tenses straight from meaning is compared with a task that is almost exactly the same: subjects produce past tenses from meaning, but the stem of the past tense is presented subliminally as a masked prime. The close correspondence of the two tasks is intended to control for variability in how predictable the target lexemes are from their contexts. Because regulars are more phonologically similar to their stems than irregulars, it is expected that priming the stem will lead to a greater difference between regulars and irregulars. If a difference between regulars and irregulars is found in this experiment, we want to be sure that it is due to the regularity status of the targets and not to variation in how well the contexts predict their targets, so the second experiment is a control for the variability of the contexts. The same target lexemes are produced from the same contexts, but in the present tense, where regularity is not a factor. The results from this experiment should provide a measure of how well the contexts predict their specific lexemes. The third experiment is a less conservative version of the form condition in the first experiment; a stem prime is given, but the subject is consciously aware of it. In the fourth experiment, the form to form transformation task of previous studies (see Section 2.3) is used with the target words of this study; the stem is shown alone, without context, and the subject produces its past tense. The results from this experiment provide a measure of how long it takes to transform a stem to past tense without the influence of context.

3.1.2 The meaning task and the form task

A crucial aspect of the proposed study that has not yet been thoroughly discussed is the operationalization of meaning in the meaning to form task. In the two meaning to form studies described above, Stemberger and MacWhinney (1988) used spontaneous speech errors as data, and Ellis and Schmidt (1998) used pictures to elicit inflected forms. Corpus data from natural speech is inappropriate for this study because the measures of interest, production reaction time and brain images, cannot be collected from existing corpora, and variables of interest would be too difficult to control if collected from spontaneous generation. Picture stimuli are a possible option, but the use of pictures presents another set of problems. The pictures used in Ellis and Schmidt (1998) represented noun concepts. Verbs tend to be more conceptually complex than nouns (Gentner, 1981) and score lower on imageability scales, so it is harder to make simple pictures that will reliably elicit specific verbs. Additionally, a form to form mapping task where the subject sees a word and must supply the past tense is perceptually much different from a task where the subject sees a picture instead. For validity of comparisons, the compared conditions should be as perceptually similar as possible.

The task I have used here is a sentence completion task where two sentences provide a coherent context that will encourage the subject to use a specific verb in a third sentence with a missing word. For example, the subject sees, *There was one more place at the conference table. He went to the chair, and he ____ down*, to which *sat* is the appropriate response. The subject sees one clause at a time, starting at the top of the screen. The stimulus requires the subject to integrate the meaning of the introductory sentences and use that coherent meaning to generate a verb in the past tense. The use of the past tense in the responses is motivated by the contexts, and not an explicit instruction. Sentences were balanced across conditions with respect to the optional object pronouns, dative prepositions and place adverbials that followed the blank. (Examples: object pronoun with dative preposition *And he ___ on it.* Place adverbial *So he ____ there.*) The verbs of interest in the study did not appear in the context setting sentences. Additionally, no regular past tense forms appeared in the context sentences. Tense was conveyed with past progressive forms (*was listening*) or with a limited set of very frequent, relatively semantically empty, irregular verbs and modals (*was, got, would, could, got, put, did, had, made, became*). None of these verbs were target words meant to be generated by the subject in the test sentences. The stimuli developed for this

Figure 3.1: Diagram of meaning task



study are shown in Appendix A.1.

The primary problem with using stimuli of this type is that context sentences will vary a lot in terms of how well they predict the target word, making it difficult to distinguish regularity effects from context predictability effects. Both accuracy and speed of target production are affected by the predictability of the target from its context (Cohen and Faulkner, 1983; Daneman and Green, 1986; Goldman-Eisler, 1958; Griffin and Bock, 1998; Lachman, 1973). Certain words may take longer to get because the context sentences preceding them simply don't make them as available as the context sentences do for other words.

Figure 2.1 was a very simple diagram of a meaning to form mapping task. The task used in this study requires a further breakdown of the 'meaning' side of the diagram as shown in Figure 3.1.

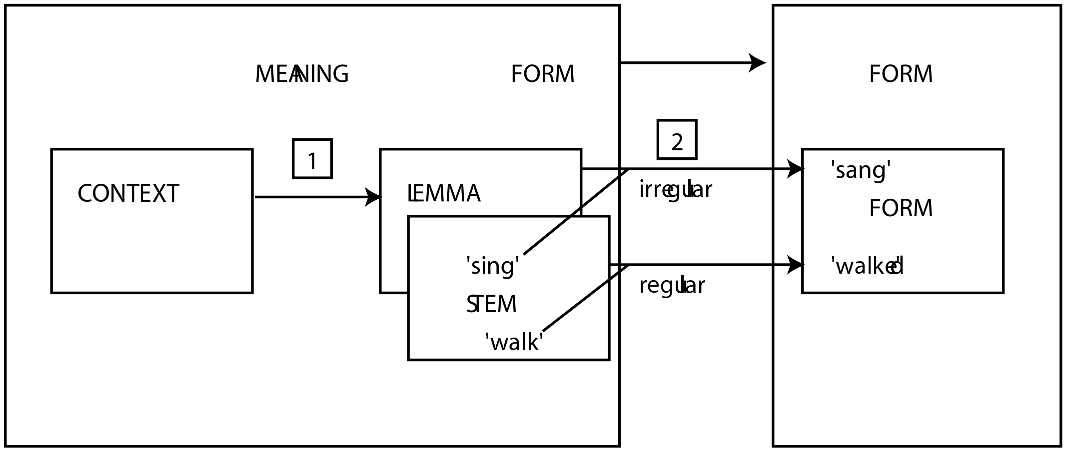
The context sentences presented at each trial lead the subject to a particular lemma at the arrow labeled $\boxed{1}$. A lemma, also referred to as a lexical entry, is the concept captured by a particular word in a language without any information about its phonological form, and most theories of language production posit a stage of lemma access. (Dell et al., 1997; Garrett, 1975; Kempen and Huijbers, 1983; Levelt, 1989; Roelofs, 1992). That lemma then takes on a phonological form at $\boxed{2}$, which is either regular or irregular. The experimental question is whether regulars and irregulars differ at $\boxed{2}$. However, if a difference is found between regulars and irregulars, it may originate in a difference in lemma predictability at $\boxed{1}$. Both $\boxed{1}$ and $\boxed{2}$ are being timed together, and it is necessary to separate out the effect of a difference in $\boxed{2}$.

This general problem of variability in the predictive power of context sentences was dealt with in a few different ways. The first method for controlling the variability of the predictive power of the contexts was built into the stimulus development phase. Originally, 124 test items were generated (test items are two context sentences followed by a fill-in-the-blank sentence). These test items were given in written survey form to 40 undergraduates at the University of Chicago. The items with the most intersubject agreement (percentage of people responding with the target word) were chosen such that there were 36 regular and 36 irregular target words. The overall intersubject agreement for the chosen items was high ($> 96\%$) and the agreement scores for regulars (96.7%) and irregulars (95.8%) did not differ significantly ($z = 1.259$, $p < .21$). Basically, these numbers say that people are very likely to respond with exactly the word they are meant to respond with in the given contexts.

Using the measure of intersubject agreement, stimuli were chosen to minimize variability in the contexts with respect to how likely a subject was to produce the target. However, this measure did not necessarily minimize differences in how *quickly* a subject would respond with the target. Because response time is the dependent variable, it is important to control for this type of variability. If one or two particular words in the regular or irregular groups take much longer to produce from their contexts, they could create a bias that looks like a regularity effect, but is actually a context effect. In order to control for this possibility, analyses were done on a by-item basis as well as on a by-subject basis where possible (Clark, 1973). The by-subject analysis captures how consistently subjects are behaving with respect to the words. The by-item analysis captures how consistent average reaction times for words are with respect to subjects. When results are significant in both analyses, it can be reasonably assumed that the classes of words (regular, irregular, high frequency, low frequency) are behaving consistently as groups, and not being thrown off by one or two particularly slow or fast members.

While the use of both by-subject and by-item analyses deals with the possibility that results are being skewed by one or two particular words, it does not deal with the possibility that the stimuli have been created in such a way that the predictability of words from their contexts is correlated with regularity or frequency. This possibility is controlled for by using the same context sentences in both the meaning and the form tasks. In the meaning task, the subject sees the two context sentences and then responds to a fill in the blank sentence. In the traditional form task, the subject sees the stem of a verb and responds with the past tense. The meaning task in the present study differs from this traditional form task because the target word must be determined from context. If there were some bias in the predictability of the words from their contexts, that bias would only affect the meaning task, and not the form task. Therefore, for the form task in the present experiment, the subject sees the same two context sentences, followed by a fill-in-the-blank sentence, but right before the fill-in-the-blank sentence, they see a brief masked prime of the stem. This task is diagrammed in Figure 3.2.

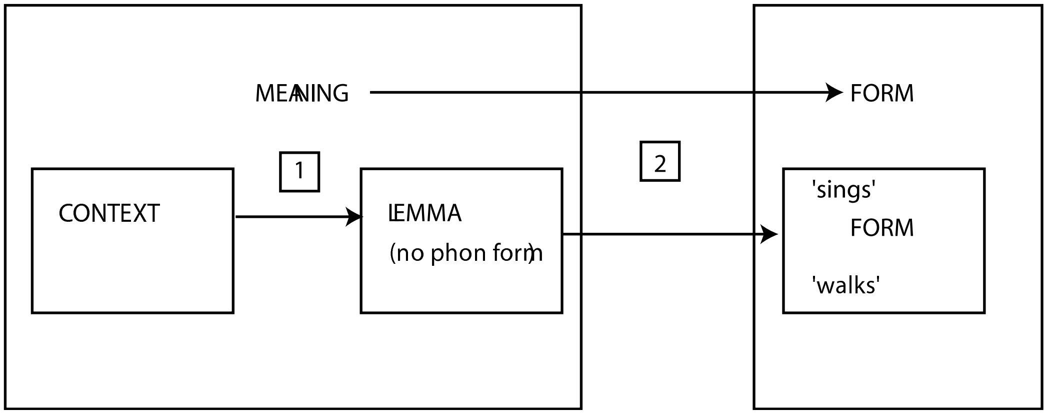
Figure 3.2: Diagram of form task



Because the same contexts are making the same words more or less accessible in both conditions, if there is a bias in the contexts, it will at least affect both of the conditions of interest equally. This reformulation of the form task means that this study does not strictly compare a form to form mapping task with a meaning to form mapping task. However, the research question of this study is only concerned with the form to form mapping task because of the influence of seeing the stem might have upon the retrieval of the past tense form. Therefore the important issue is whether the subject is exposed to the stem or not. The form task used here is not exactly a form to form mapping task. Instead, it is a meaning to form mapping task that includes stem exposure. The main theoretical question is preserved while making the conditions to be compared as similar as possible.

Another method of controlling for context variability was to run a test with a different set of subjects, in which the task was to produce the present tense of the verb. In this condition, instead of seeing *There was one more place at the conference table. He went to the chair, and he ____ down*, the subject sees *There is one more place at the conference table. He goes to the chair, and he ____ down.* The target

Figure 3.3: Diagram of present tense task



word in this case is *sits*. There are no regularity differences in the present tense, because all present tense forms are regular.

This task does not mix context predictability effects with regularity effects because there is no regularity effect for the present tense. As shown in Figure 3.3, it measures the time it takes to come up with the lemma from the context ($\boxed{1}$) plus a constant ($\boxed{2}$). This task served as a baseline measurement of reaction time to target words in their contexts.

3.1.3 Stimulus characteristics

The choice of past tense forms to be generated was constrained in two ways by the development of the meaning to form task. First, not every word can be adequately constrained by context this task. It can be very difficult, for example, to come up with a context that will make a person likely to say *began* but not *started* or *sought* but not *looked for*. For some very low frequency words, it is difficult to come up with any context that would predict it, as is the case with *trod*. Second, the items selected from the original larger test set of 124 had to meet a high level of intersubject agreement. Luckily, the top scoring eighty or so items were well distributed in that half were regular and half were irregular. Furthermore, they were well distributed in terms of frequency. Verbs were labeled as high frequency if their past tense forms had a frequency of above 35 per million according to Francis and Kučera (1983) and as low frequency if below 35. (This was the cutoff point between low and high frequency used in Stemberger and MacWhinney, 1988). The frequency characteristics of the verbs are summarized in Appendix A.2. The stimulus set was balanced so that there were no significant differences between the low frequency regulars and irregulars, or between the high frequency regulars and irregulars, with respect to absolute past tense frequency, present tense frequency, or verb cluster (the verb in all its inflected forms) frequency. In other words, regularity does not correlate with any frequency measures.

3.2 Experiment 1: Meaning vs. Subliminal Form

This experiment investigated the difference between generating a past tense form from meaning with and without exposure to the verb stem form. It tests the hypothesis that exposure to the stem form makes regulars and irregulars differ. In the meaning condition (no stem form present), subjects responded to a past tense context followed by a fill-in-the-blank sentence with a word appropriate to complete the sentence meaningfully. In the subliminal form condition (stem form present), subjects responded to the same stimuli as in the meaning condition, except that the stem of the intended past tense target was displayed for a very brief instant right before the fill-in-the-blank sentence. If stem exposure causes regulars and irregulars to differ, irregulars should have different RTs than regulars in the subliminal form condition, but not in the meaning condition.

3.2.1 Methods

Participants

A total of 42 students at the University of Chicago were run in both conditions together (meaning 21, subliminal form 21). They were paid for their participation.

Materials

The set of stimuli has been described above in Section 3.1.2. In the meaning condition, no stem form was shown. In the subliminal form condition the stem form was displayed as a masked prime.

A practice list of 5 test items not included in the experimental list was also created.

Procedure

Stimuli were presented in random order using E-Prime (2000) software. Participants were told that they would see context sentences followed by a fill-in-the-blank sentence, and they should speak the word that belonged in the blank into a microphone. They were told to say the first word that occurred to them. They were told that the microphone registered how long it took them to respond and they should respond as quickly as possible. The phrase *past tense* was not used in the instructions. They were simply told to speak the most obvious word for the context.

The visual display space was divided into three equal horizontal frames, and the three sentences of each test item were presented in order from top to bottom in the center of each frame. The first context sentence was presented for 2000 ms before the next sentence appeared. The presentation time of the second context sentence varied according to how long the sentence was. It was presented for 120 ms multiplied by the number of syllables in the sentence. [1] Times ranged from 720 to 1920 ms and there were no significant differences in length of second context sentence for any variables of interest, as shown in Appendix A.2. Response times were recorded from onset of the presentation of the third sentence, and collected with E-Prime's voice key feature. In the subliminal form condition, the third sentence was preceded by a 60 ms presentation of the stem [2] in third sentence position, followed by a 200 ms mask. This slight delay between the second context sentence and the fill-in-the-blank sentence was present in the meaning condition as well, but in that case a 60 ms presentation of one mask was followed by a 200 ms presentation of another mask. The sentences disappeared as soon as the subject responded, or if the subject did not respond, after 6000 ms, and were replaced by a fixation point at the top of the screen that remained for 2000 ms until the next item began. The 72 test items were presented after the 5 item practice session. A schematic of the time course and types of events for each condition is shown in Figure 3.4.

The experimenter kept track of subjects' responses by marking a list of target words, indicating whether they said the correct word, said a different lexical item, made a past tense error (i.e., overregularization), or had a voice key error.

3.2.2 Results

Response time data were included only for correct responses, and three subjects were excluded from subsequent analysis due to error rates of over 30%. Responses

1. This is at the upper end of average reading speed as calculated from data in Dellwo and Wagner (2003).

2. This is the point at which improvements in masked priming effects reach asymptote as reported in Forster (1999).

Figure 3.4: Procedure for experiment 1

	Meaning	Subliminal Form
Fixation 2000 ms	+	+
Clause 1 2000 ms	There was one more place at the conference table.	There was one more place at the conference table.
Clause 2 720-1920 ms	He went to the chair.	He went to the chair.
Prime 60 ms	XXXXX	SIT
Mask 200 ms	++++++	++++++
Clause 3 until response or 6000 ms	and he ____ down.	and he ____ down.

were also excluded from analysis if the RT was greater than 2000 ms or less than 200 ms. An additional subject was excluded for having over 30% of responses over 2000 ms. After error exclusion and trimming 38 subjects remained (meaning 20, subliminal form 18). Of the responses of the remaining subjects 16% were trimmed. This proportion of eliminated responses is normal for sentence frame elicitation paradigms (see Griffin (2002), where 18-25% of responses are eliminated in each experiment). There were no statistical differences, with respect to the proportion eliminated, between regulars and irregulars, or between high and low frequency words as measured by t tests. During the trimming process, I noticed that some words seemed to have unusually long RTs (and thus were often being trimmed). A by-items analysis revealed that the RTs for three words were more than two standard deviations from the mean of all words: *bleed*, *freeze*, and *shed*. These words all shared something in common; their fill-in-the-blank sentences began with a full noun phrase rather than a pronoun. These were the only three test items that did not start the fill-in-the-blank sentence with a pronoun. They were removed from the

Table 3.1: Means(SD) for experiment 1

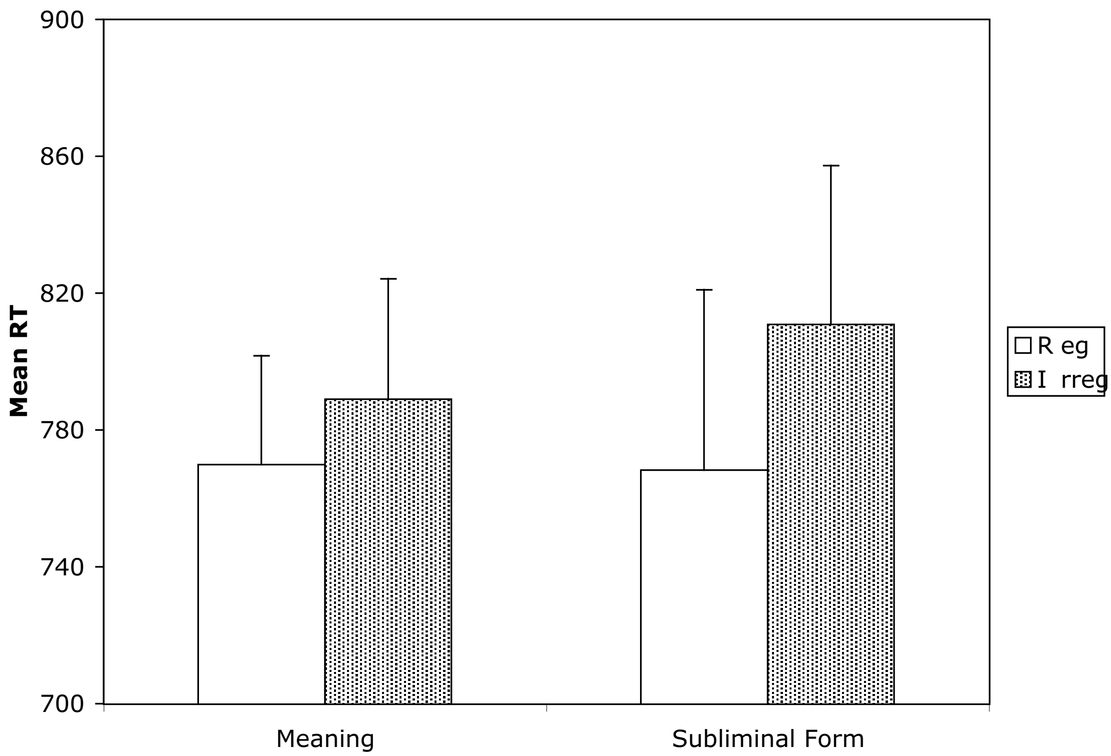
Meaning				Subliminal Form			
Regular		Irregular		Regular		Irregular	
770(142)		789(158)		768(224)		811(197)	
Hi	Lo	Hi	Lo	Hi	Lo	Hi	Lo
764(159)	772(148)	793(189)	783(146)	748(214)	786(257)	804(185)	816(217)

analysis, leaving a total of 69 words and subject averages were recalculated. It was also confirmed that the removal of these items did not cause frequency measures or duration of the second context sentence to differ significantly between any variables of interest (see Appendix A.3).

A 2 x 2 repeated measures ANOVA with regularity and frequency as within subjects factors was run on each condition. A marginally significant main effect of regularity where irregular verbs took longer than regular verbs by 43 ms was found in the subliminal form condition both by subject $F(1, 17) = 4.067$, $p < .06$ and by item $F(1, 65) = 3.371$, $p = .07$, but regularity had no main effect in the meaning condition by subject or by item ($F < 1$). No main effect of frequency was found for either condition ($F < 1.4$). The Regularity X Frequency interaction had no effect in meaning, or subliminal form ($F < 1$). The means and standard deviations are reported in Table 3.1.

The regularity effect was greater in the subliminal form condition (43 ms) than in the meaning condition (19 ms), but perhaps not significantly greater. In order to see whether the two conditions differed significantly from each other, the regularity effect for each subject was computed, and a t test between the conditions was performed. The subliminal form condition did not differ significantly from the meaning condition $t(36) = -.759$, $p = .45$. The regularity effects for the two conditions are depicted in Figure 3.5.

Figure 3.5: Mean RTs for irregular and regular in meaning and subliminal form conditions.



The initial analyses revealed no significant difference in the time it takes to produce regular and irregular past tense verbs when those verbs are retrieved straight from a meaningful context. In those same meaningful contexts, the regular past tense verbs are produced more quickly than the irregulars when the subject has been primed by the stem of the verb, though the difference is only marginally significant.

3.2.3 Discussion

There was a marginally significant effect of regularity, where irregulars had higher reaction times than regulars, in the subliminal form condition, but not in the meaning condition. There was no main effect of frequency found, nor was the expected Frequency X Regularity interaction found. It is not too surprising that frequency failed to play a significant role in this experiment. In a task where context drives the choice of words, frequency should be of less importance than that context in determining how quickly a word is retrieved. Frequency influences word retrieval in a default situation, but when a context is given, even a generally infrequent word can be highly predictable within that context, and thus be more quickly retrieved. For example, the word *shed* is low frequency, but in the context evoked by *dogs* and *hair* it is probably easier to retrieve than in a neutral context. Indeed a study by Griffin and Bock (1998) elicited words using pictures that followed sentences of higher or lower levels of contextual constraint. When the pictures were shown in isolation, low frequency words were much slower to produce than high frequency words, but the frequency effect disappeared at high levels of contextual constraint. Subjects were faster to produce a low frequency word like *pumpkin* after seeing *For Halloween, they carved up a large. . .* Grosjean and Itzler (1984) also found that frequency effects disappeared with highly constraining contextual support in a gating study (where the amount of a word displayed is increased until it is identified).

There was a marginally significant effect of regularity in the subliminal form condition, but not in the meaning condition. This is consistent with the idea that exposure to the stem form has an effect on the production of the past tense form. The results are not consistent with an account where irregulars are qualitatively different from regulars when they are produced straight from meaning, without stem exposure. However, when the conditions were compared directly to each other, no difference in the regularity effect between the two conditions was found. Although the effect of regularity was not significant in the meaning condition, the mean for irregulars is longer than that for regulars by 19 ms. It may be the case that the meaningful contexts contribute a slight bias that makes irregulars have longer RTs. The difference is enhanced in the subliminal form condition, but because it is also there in the meaning condition, the comparison of the conditions to each other doesn't reveal a difference. A follow up experiment was conducted to see whether the contexts led to what looked like a regularity effect.

3.3 Experiment 2: Present tense as context measure

Experiment 1 found that irregulars appear to take longer to produce than regulars when a stem prime was given, but not when no prime was given. However, there was no significant difference between the prime and no prime conditions when compared to each other. It is possible that differences in how well the contexts predict their target lemmas result in response time differences that are in the same direction as the regularity effect (thus making differences in the regularity effect between the two conditions difficult to detect). In this experiment I determine whether the contexts used in Experiment 1 are biased in a way that make irregulars have longer RTs. This experiment used the same contexts and target words as in Experiment 1 except the contexts were in the present tense and the target word was meant to be in the present tense. There is no regularity difference in the present tense, so response times in this task should reflect how accessible the target lemmas are in those contexts, and any difference between regulars and irregulars should reflect differences in context predictability and not regularity effects.

3.3.1 Methods

Participants

20 students at the University of Chicago participated in this experiment. They were paid for their participation.

Materials

The same set of stimuli as used in Experiment 1 were used here, except context sentences occurred in the present tense. Instead of seeing *There was one more place at the conference table. He went to the chair, and he ____ down*, the subject saw *There is one more place at the conference table. He goes to the chair, and he ____ down.* The target word in this case was *sits*.

Procedure

The procedure was the same as the meaning condition in Experiment 1, but with present tense context sentences. An example is given in Figure 3.6.

Figure 3.6: Procedure for experiment 2

	Present
Fixation 2000 ms	+
Clause 1 2000 ms	There is one more place at the conference table.
Clause 2 720-1920 ms	He goes to the chair.
Prime 60 ms	XXXXX
Mask 200 ms	++++++
Clause 3 until response or 6000 ms	and he ____ down.

3.3.2 Results and discussion

Reaction time data were included only for correct responses between 200 and 2000 ms. 17% of responses were eliminated. Again this proportion of eliminated responses is normal for sentence frame elicitation paradigms and t tests confirmed that there were no statistical differences, with respect to the proportion eliminated, between regulars and irregulars, or between high and low frequency words. One subject was excluded because he responded with the past tense instead of the present tense, leaving a total of 19 subjects.

The present tense condition was intended to serve as a measure of how difficult it is to access the target lemmas in their respective contexts. There should be no effect of regularity in the present tense condition because present tense verbs are all regular. Indeed, there was no significant effect of regularity in the present condition, nor was there an effect of frequency or Regularity X Frequency interaction. Means and standard deviations are reported in Table 3.2. No main effects or interactions

Table 3.2: Means(SD) for experiment 2

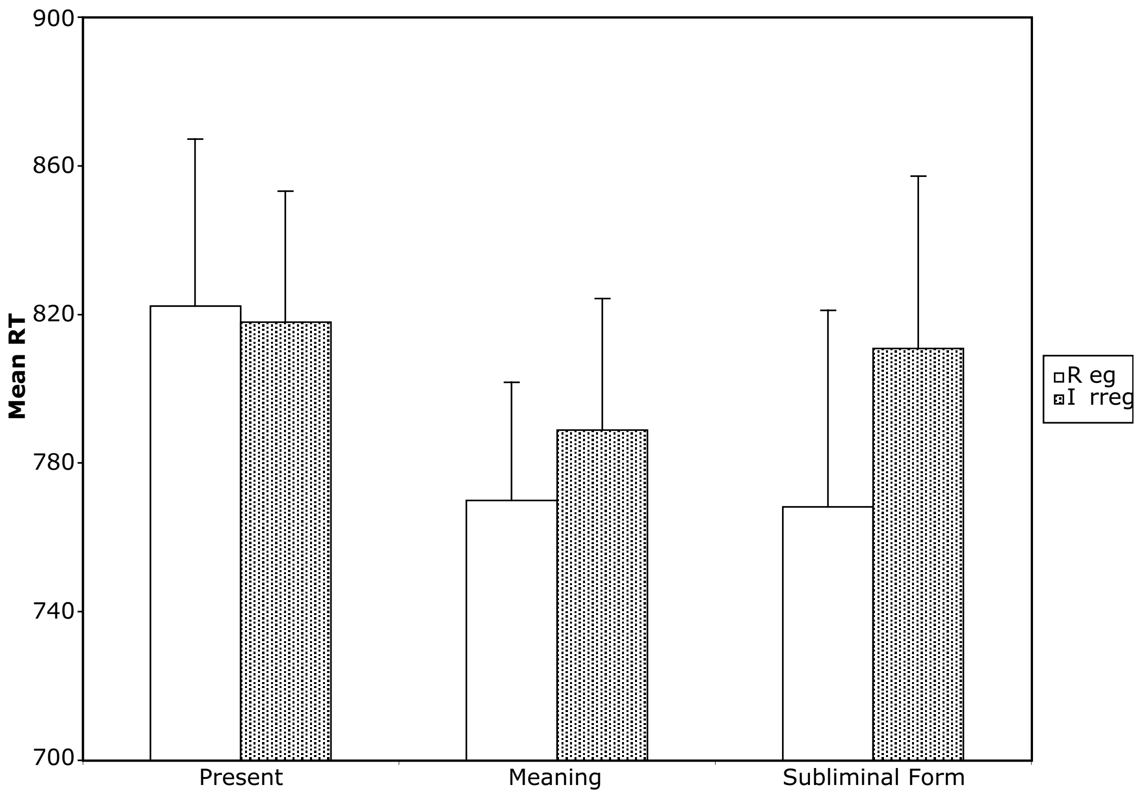
Present			
Regular		Irregular	
822(190)		818(150)	
Hi	Lo	Hi	Lo
807(202)	830(192)	825(167)	808(133)

approached significance. Additionally, the mean difference between irregulars and regulars was -4 ms. There appears not to be a bias in the contexts making the irregulars take longer.

A t test comparing the regularity effect in the present condition with the meaning condition of Experiment 1 did not reveal a significant difference between the two conditions, $t(36) = .811$, $p = .42$. However when the present tense was compared to the subliminal form condition, there was a marginally significant difference between the two conditions, $t(34) = -1.681$, $p = .10$. The differences in the regularity effect between the conditions is shown in Figure 3.7.

There is no difference between regulars and irregulars in the present tense condition, indicating that contexts do not create a bias against irregulars or for regulars. There is a slight, non-significant difference of 19 ms between regulars and irregulars in the meaning condition, but this regularity effect is not statistically greater than the -4 ms effect of the present tense. There is also no difference between the regularity effects of the meaning and the subliminal form conditions. However, there is a marginal effect of regularity in the subliminal form condition, and a marginally significant difference in the regularity effect between the subliminal form and the present condition. In summary, with respect to regularity, generating the present tense is no different from generating the past tense *when the subject has not seen the stem*. When the subject *has* seen the stem, there *is* a difference between generating

Figure 3.7: Mean RTs for irregular and regular in present, meaning, and subliminal form conditions.



the present and the past tense-the irregular verbs of the past tense take longer to produce. This summary carries the caveat that results of 'difference' and 'longer to produce' attain only marginal statistical significance.

The RTs collected in the present tense condition are a measure of how long it takes to come up with the target lemmas in their contexts. Analysis of these times show that, as a group, the words which would be irregular in the past tense take no longer to come up with from their contexts than the regular words do. We can also use these times for analysis not just of groups, or categories, of words, but also in an analysis where we consider individual measures of how long it takes to come up with each word in its context. We can regress those measures on the measures for the other conditions and see how well they correlate. It could be the case that the time required to come up with the target words from their contexts is accounting for much of the variation in all of the conditions and making it harder to detect differences between the conditions with respect to regularity. Average RTs for each word in the present condition were regressed on the average times from the other conditions in order to see whether context accounted for much of the variance, and if so how much.

The average reaction times by item for each condition were regressed on the present tense times. The regression was significant for both of the conditions (subliminal form $F(1, 67) = 17.27$, $p < .0001$, meaning $F(1, 67) = 46.117$, $p < .0001$). R^2 was highest for the meaning condition ($R^2 = .41$), followed by the subliminal form condition ($R^2 = .21$). This says that the difference between meaningful contexts accounted for 41% of the variation in the meaning condition, but for only 21% of the variation in the subliminal form condition. This result will be returned to in Section 3.4.2.

So far, nothing indicates that there is a difference between regulars and irregulars in the present tense or meaning conditions. There is some indication that they are different in the subliminal form condition, but only through marginally significant results.

It could be the case that the subliminal form condition was too conservative. The present study argues that the activation of the phonological form of the stem has an effect on the ability to activate the phonological form of the past tense. In this subliminal form task, the subjects appear to be primed by the stem (as indicated by the marginally significant result), but it is not necessarily the case that their phonological forms are activated. In the literature on masked priming, there is general agreement that semantic priming does occur with masked primes, but a lack of consensus on whether phonological priming occurs with masked primes. Some studies find priming effects (Ferrand and Grainger, 1992; Ferrand and Grainger, 1994; Lukatela, Frost, and Turvey, 1998), and some do not (Davis, Castles, and Iakovidis, 1998; Shen and Forster, 1999) even at similar prime durations. One study approached these inconsistent results by controlling for conscious awareness of the prime. Kouider, Peereman, and Dupoux (), in a study manipulating prime duration found that when subjects were divided into groups based on whether they had a conscious awareness of the prime or not, it was conscious awareness, regardless of prime duration, that led to phonological priming effects.

In the present study, subjects in the subliminal form condition had been asked after the experiment whether they noticed anything happening right before they saw the final sentence. Nine subjects reported that they saw nothing or “a program glitch” and 9 reported that they saw a word, either “the word I’m supposed to say” or “a word related to the one I’m supposed to say.” The results were separated into two groups, one of subjects who noticed the prime and one of subjects who did not, and repeated measures ANOVAs with regularity as a within subject factor were run on each group. For the group who noticed the primes, regularity was significant $F(1,8) = 11.667$, $p < .01$, but it was not significant for those who did not notice the prime ($F < 1$). In order to make the activation of the phonological stem more robust, an additional experiment was run. In this experiment, instead of the conservative masked form priming, the stem form was fully displayed inside the blank.

3.4 Experiment 3: Stem prime with conscious awareness

In Experiment 1 a subliminal form prime resulted in a difference between regular and irregular verbs, but this difference was only marginally significant. Additionally, the difference was not significantly greater than that found when no prime was given. It is possible that for stem exposure to have an effect on the production of past tense verbs, the subject must have conscious awareness of the stem presentation. This experiment used a form task where the subject read the stem form before responding with the past tense form. However, in order to be comparable to the meaning task, this stem form was read within the same meaningful contexts used in the meaning condition. It differed from the subliminal form task in that all subjects were consciously aware of the stem form being presented, and therefore its phonological form was more likely to be activated before the past tense form was produced. This task will be referred to as the form condition.

3.4.1 Methods

Participants

15 students at the University of Chicago were run in this experiment. They were paid for their participation.

Materials

The same set of stimuli used in Experiment 1 were used here, and an additional 20 test items were added. In this condition, the stem of the target past tense word is clearly given to the subject. Because of this, the subject could adopt a strategy of just waiting for the word in the blank to appear and saying the past tense of that word without reading the preceding context sentences. In order to keep the subject from adopting this strategy and ensure that the context sentences would be read, the additional 20 items had no stem in the blank position. Therefore, at the beginning of a trial, the subject did not know whether a stem form would be given, and so it was necessary to read the context sentences in every case.

Procedure

Participants were told that they would see context sentences followed by a fill-in-the-blank sentence, and they should speak the word that belonged in the blank into a microphone. There were told that sometimes a word would be in the blank, but it might not be in the correct form for the context, and in that case they should say the correct form of that word. There was no mention of 'past tense'.

Other aspects of the procedure were the same as in Experiment 1. In order to be consistent with the conditions in Experiment 1, before the fill-in-the-blank sentence

Figure 3.8: Procedure for experiment 3

	Form
Fixation 2000 ms	+
Clause 1 2000 ms	There was one more place at the conference table.
Clause 2 720-1920 ms	He went to the chair.
Prime 60 ms	XXXXX
Mask 200 ms	++++++
Clause 3 until response or 6000 ms	and he _SIT_ down.

there was a 60 ms presentation of one mask followed by a 200 ms presentation of another mask. The procedure is shown in Figure 3.8.

3.4.2 Results and discussion

Again, reaction time data were included only for correct responses between 200 and 2000 ms (6% eliminated, equal proportions of regular and irregular and of high and low frequency). A 2 x 2 repeated measures ANOVA with regularity and frequency as within subjects factors revealed a main effect of regularity by subject $F(1,14) = 37.006$, $p < .0001$ and by item $F(1,65) = 13.064$, $p < .001$. No main effect was found for frequency, but there was a Frequency X Regularity interaction $F(1,14) = 13.234$, $p < .005$ that was marginal by item $F(1,65)3.893$, $p < .06$. Means and standard deviations are given in Table 3.3.

This form condition was compared to the conditions from the previous experiments using t tests of the regularity effect (difference between regulars and irregu-

Table 3.3: Means(SD) for experiment 3

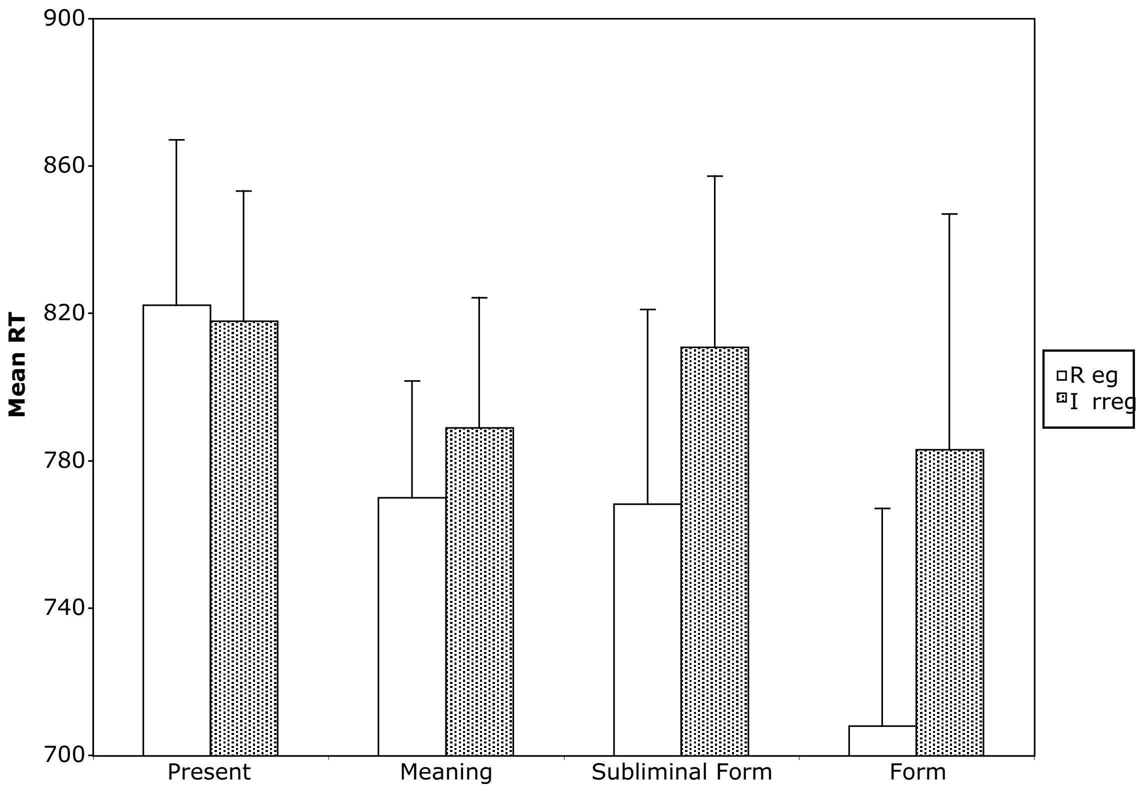
Form			
Regular		Irregular	
708(227)		783(247)	
Hi	Lo	Hi	Lo
707(209)	725(244)	816(262)	752(234)

lars) by subject. The regularity effect for this form condition was 75 ms which did not differ from the 43 ms effect for the subliminal form condition, but it was significantly greater than the (-4 ms) present condition effect $t(31) = -3.415$, $p < .002$ and greater than the (19 ms) meaning condition effect to a marginally significant degree $t(33) = -1.988$, $p < .06$. Irregulars are slower to produce than regulars in the form condition. Irregulars and regulars are not different in the meaning condition. With respect to the difference between regulars and irregulars, the meaning condition does not differ from the present condition. The form condition does differ from the present condition, and from the meaning condition. The mean regularity effects for the conditions are shown in Figure 3.9.

This experiment used a task that activated the phonological form of the stem to a greater degree than the subliminal form condition in Experiment 1. As was the case with the subliminal form condition, an effect of regularity was found, with irregulars having longer reaction times than regulars, but for the form condition in this experiment, the result was unambiguously significant. Additionally, the regularity effect of 75 ms in this condition was significantly greater than the effect for the present condition and marginally significantly greater than that for the meaning condition.

How much variation is explained by context in this condition? The regression of form RTs on the RTs for the present condition was significant $F(1, 67) = 9.844$,

Figure 3.9: Mean RTs for irregular and regular in present, meaning, subliminal form, and form conditions.



$p < .003$, $R^2 = .13$. Context accounts for a significant part of the variation in this condition, but less than in the other conditions. The difference between meaningful contexts accounted for 41% the variation in the meaning condition, for 21% of the variation in the subliminal form context, and for 13% of the variation in the form condition.

The regression analysis revealed a characterization of the data where the effects of meaningful context contribute the most in the meaning condition, and the least in the form condition. It may be the case that the remaining variation can be accounted for by the transformation of the stem form into the past form, but this cannot be determined without a measurement of this transformation effect. A measure of this transformation effect would be the reaction times to the simple version of the form to form mapping task where the subject sees the stem form, and responds with the past form, without any meaningful context. The present condition was a 'pure' meaning task; it yielded a measurement of the contribution of context to the production of a word apart from regularity or form transformation effects. The opposite end of the scale from this condition would be a 'pure' form task; it would yield a measurement of the contribution of the transformation effect to the production of those same words. A follow up experiment was conducted in order to assess the contribution of the transformation effect to the other conditions.

3.5 Experiment 4: Simple stem to past mapping task

The form to form task used by previous studies is assumed to reveal facts about regularity. While it is unclear whether the results of those studies do reveal facts about regularity, or whether they only reveal facts about the form to form mapping task, it is clear that they measure the ability of subjects to make a direct transformation between the stem form and the past tense form. Because this ability is assumed to reveal facts about the effects of regularity in natural language, it can be considered a covert assumption that a transformation from stem to past is involved in everyday situations where the past tense is produced. In the previous experiments, the four conditions range from a pure measure of context predictability (present condition) to a measure of context predictability combined with the form transformation (form condition). This follow up experiment completed the continuum by adding the other end of the scale: pure form to form transformation. The same past tense words were produced as in the previous experiment, but instead of being prompted by context, or by a combination of context and stem form, they were prompted by a simple stem form.

3.5.1 Methods

Participants

A total of 16 students at the University of Chicago participated in the experiment. They were paid for their participation.

Materials

The same set of stimuli was used as in the other experiments. However, the context sentences were not used. Subjects saw the stem of a word, and responded with the past tense.

Figure 3.10: Procedure for experiment 4

	Simple
Fixation 2000 ms	+
Stem until response or 5000 ms	SIT

Procedure

Again, stimuli were presented in random order with E-prime software. Subjects were told that they would see a verb in the middle of the screen, and they should say the past tense of that verb into a microphone.

Each trial began with the 2000 ms fixation point in the center of the monitor. Then the stem was presented in all capitals in the center of the monitor. The procedure is shown in Figure 3.10. Reaction times were collected by E-prime's voice key feature. The stem disappeared as soon as the subject responded, and the next trial began. Subjects had 5000 ms in which to respond. The 72 test items were presented after 5 practice trials.

The experimenter kept track of subjects' responses by marking a list of target words, indicating whether they said the correct word, said a different lexical item, made a past tense error (i.e. overregularization), or had a voice key error.

3.5.2 Results and discussion

Data from one subject were excluded because of voice key errors. Reaction time data were included only for correct responses between 200 and 2000 ms (6% eliminated, equal percentages of regular and irregular, high and low frequency). The three items

Table 3.4: Means(SD) for experiment 4

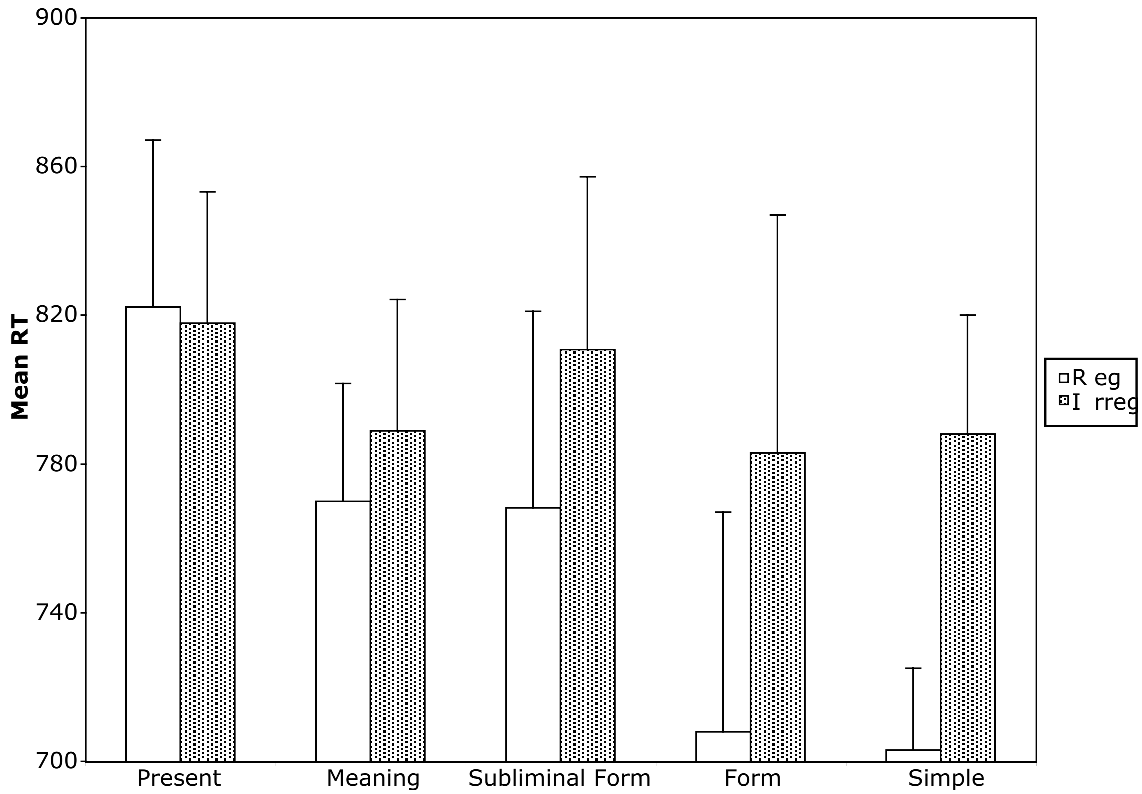
Simple			
Regular		Irregular	
703(84)		788(122)	
Hi	Lo	Hi	Lo
685(98)	713(84)	805(134)	771(115)

excluded from the analysis of Experiment 1 were also excluded from the analysis of this experiment.

This stem to form transformation task will be referred to as the simple condition. A 2 x 2 repeated measures ANOVA with regularity and frequency as within subjects factors was run. The effect of regularity was significant both by subject $F(1, 14) = 45.431$, $p < .0001$ and by item $F(1, 67) = 35.271$, $p < .0001$. The effect of frequency was not significant by subject or item ($F < 1$). The effect of the Frequency X Regularity interaction was significant by subject $F(1, 14) = 12.219$, $p < .004$ and marginally significant by item $F(1, 67) = 3.966$, $p < .06$. However, as in the other experiments, the interaction was not the expected one, where low frequency irregulars take longer than high frequency irregulars. Instead high frequency irregulars take longer than low frequency irregulars. The question of frequency will be dealt with in Section 3.6.1. Means and standard deviations are given in Table 3.4.

The Regularity Effect for this condition was 85 ms. Results from t tests of the regularity effect for this condition with the other conditions revealed a significant difference between this simple condition and the present condition $t(31) = -3.873$, $p < .0005$, and between simple and meaning $t(33) = -2.324$, $p = .0264$. It was not different from subliminal form $t(31) = 1.574$, $p = .13$ or from form $t(28) = .451$, $p = .66$. All five conditions are shown together in Figure 3.11.

Figure 3.11: Mean RTs for irregular and regular in present, meaning, subliminal form, form, and simple conditions.



In Section 3.4.2 results of a regression analysis that showed that context effects accounted for a large part of the variation in the meaning condition, less in the subliminal form condition and the least in the form condition. Experiment 4 yielded a measure of the stem to past transformation effect. The average times for each word in this task were used as a regressor on the results from the other experiments in order to discover how much variation in each of those conditions was accounted for by the transformation effect. The average reaction times for each condition were regressed on the average times for each word on the simple stem to past task. The regression was not significant for the present or meaning conditions. It was marginally significant for the subliminal form condition $F(1, 67) = 3.707$, $p < .06$, and it was significant for the form condition $F(1, 67) = 5.597$, $p < .03$. R^2 was highest for the form condition ($R^2 = .08$), followed by the subliminal form condition ($R^2 = .05$), and R^2 was 0 for the present and meaning conditions. The transformation effect accounted for none of the variation in the present or meaning conditions, for 5% in the subliminal form condition, and for 9% of the variation in the form condition. The fact that the transformation effect does not account for any variation in the meaning condition supports the claim that there is no transforming of stem to past tense involved in generation of the past tense from meaning.

In terms of proportion of variance accounted for, this pattern of results is the opposite of that found in the regression of present tense times on the other conditions. While the proportion of variance accounted for by context effects from highest to lowest is meaning, subliminal form, form, the proportion accounted for by transformation effects is form, subliminal form, meaning.

Table 3.5: Summary of results for all five conditions

	Present	Meaning	Subliminal Form	Form	Simple
Irreg-Reg difference	-4 ms	19 ms	43 ms	75 ms	85 ms
P value for Reg vs. Irreg	.81	.40	.0598	$< .0001$	$< .0001$
R^2 for context	N/A	.41	.21	.13	N/A
R^2 for transform	N/A	.00	.05	.08	N/A

3.6 Joint analysis of experiments

A summary of the results over all the conditions is given in Table 3.5.

The conditions run in the four experiments fall along a continuum in a variety of respects. The present tense condition is the extreme meaning-based condition, where all variation should be explained by the ability of the context to predict its word, and none should be explained by regularity, since there is no regularity difference in the present tense. The simple condition is the extreme form-based condition, where most variation should be explained by the ability to transform a word into its past tense. The other three conditions fall between the extremes. For the meaning condition, variation will be explained by the ability to produce words in their contexts, and also, possibly by regularity of those words. For the subliminal form condition, variation will be explained by the ability to produce words in their contexts, but also possibly by regularity, and by the exposure to a stem form. For the form condition, variation will be explained by the ability to produce words in their contexts, by greater exposure to a stem form, and by regularity.

The results are consistent with this view of the continuum. Context explains the most in the meaning condition, and the least in the form condition. Form transformation explains the most in the form condition and none in the meaning condition.

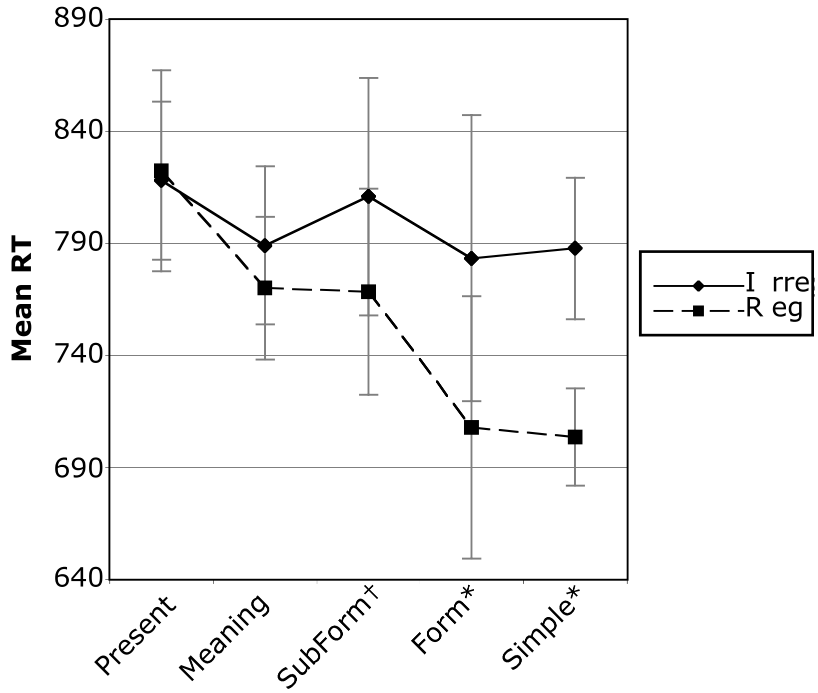
The regularity effect is not significant in the present or meaning conditions. It is marginally significant in the subliminal form condition, and significant in the form and simple conditions, and additionally, the probability that there is no difference between regulars and irregulars is less and less, going across the continuum (as reflected in the p values).

Figure 3.12 shows a graph of the means for irregular and regular in each condition as well as a table indicating which conditions differ from each other by a t test on their regularity effects.

The graph illustrates, moving from left to right, how the distance between regular and irregular increases. It also reveals something more specific about the way stem exposure affects regulars and irregulars. This study has established that the distance between regular and irregular verbs is increased by stem exposure, but so far we do not know whether it does so by making regulars faster, by making irregulars slower, or both. The slope of the lines in the graph in Figure 3.12 indicates that stem exposure increases the difference between regulars and irregulars by making regulars faster. From one condition to the next, the RTs for irregulars stay basically steady, while there is a steep drop off in the RTs for the regulars.

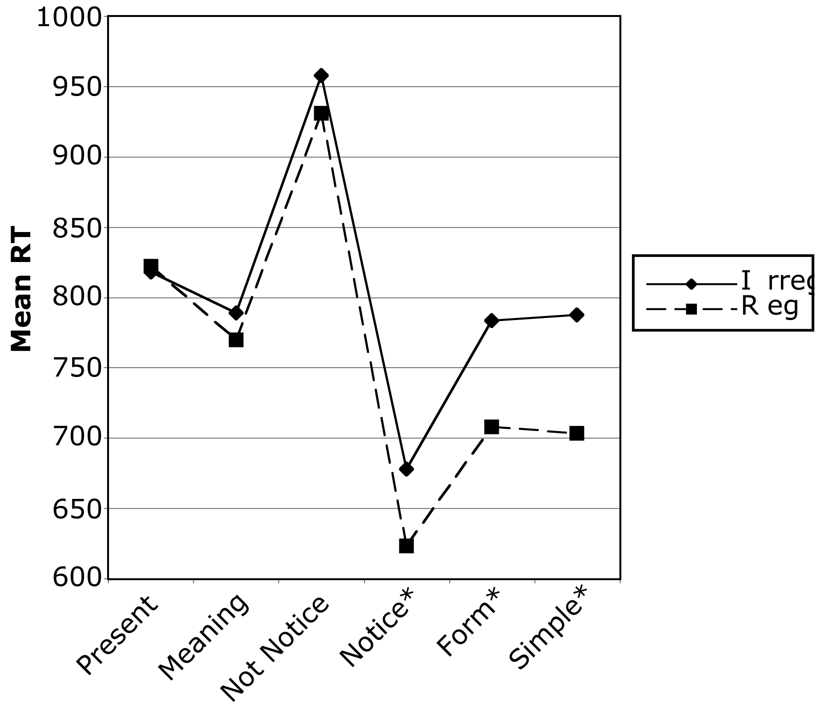
The subliminal form results keep the trend from looking smooth. These results are strange in that they represent two groups of subjects, one of which had a regularity effect—the group who noticed the prime, and one which did not have a regularity effect—the group who did not notice the prime. Splitting the subliminal form condition makes comparison with the other conditions less valid because the groups only have half the subjects of the other conditions, but it can still be informative to examine the behavior of these groups separately. As seen in Figure 3.13, the notice and not notice groups do not reach a mean reaction time consistent

Figure 3.12: Mean RTs for irregular and regular in present, meaning, subliminal form, form, and simple conditions. The symbol ∗ indicates a significant difference, and † indicates a marginally significant difference.



	Pres	Mng	SubForm	Form	Simple
Pres	–		†	∗	∗
Mng		–		†	∗
SubForm	†		–		
Form	∗	†		–	
Simple	∗	∗			–

Figure 3.13: Mean RTs for irregular and regular in all conditions, with subliminal form split into notice and not notice groups



with the other conditions. However, in other measures, the group who did notice patterns like the form condition and the group who did not notice patterns like the meaning condition, as seen in Table 3.6.

The regularity effect for the group who did not notice was 31 ms and that for the notice group was 55 ms. While the regularity effect for the group who did not notice did not differ from the present condition (as was the case for the meaning condition), the effect for the group who did notice was significantly greater than that for the present condition (as was the case for the form condition) $t(25) = -2.126, p < .05$. It was not, however, greater than the effect for the meaning condition, but again, there were fewer subjects.

Additionally, a regression of present tense times on the notice and not notice

Table 3.6: Summary of results for all five conditions

	Present	Meaning	Not Notice	Notice	Form	Simple
Irreg-Reg difference	-4 ms	19 ms	31 ms	55 ms	75 ms	85 ms
P value for Reg vs. Irreg	.81	.40	.48	$< .01$	$< .0001$	$< .0001$
R^2 for context	N/A	.41	.18	.15	.13	N/A
R^2 for transform	N/A	.00	.05	.01	.08	N/A

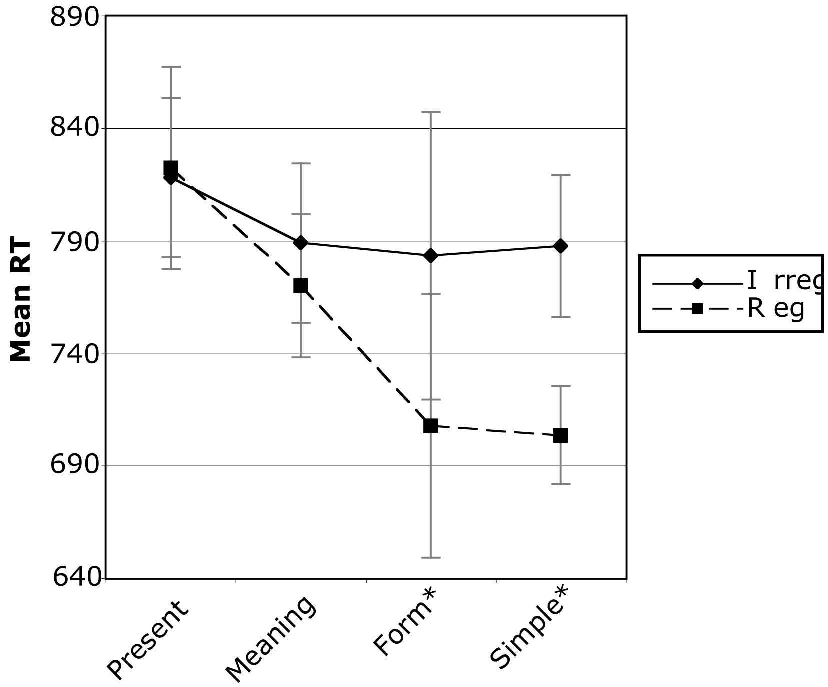
groups, found that the amount of variation explained by context for the group who did not notice was slightly higher than the amount for the group who did notice (not notice $R^2 = .18$, notice $R^2 = .15$). Again, the group who did not notice behaves closer to the meaning condition and the notice group closer to the form condition.

The trend is reversed for the amount of variation explained by the transformation effect, with the not notice group having a higher R^2 (.05) than the notice group (.01), but those R^2 values don't have any real meaning because the overall regression of simple condition times on those conditions was not significant in either case.

A post hoc investigation of the results from the subliminal form condition reveal an interesting tendency in how they pattern with respect to the other conditions, but they are perhaps less informative than desirable because they group two different types of subjects together. A clearer picture emerges from looking at the other conditions apart from subliminal form. In Figure 3.14, the trend that supports the idea that regulars are made faster by stem exposure is quite prominent. Irregulars take about the same time to produce whether a stem is shown or not, but regulars are faster when the stem is shown.

Overall, the results show that the difference between regulars and irregulars is enhanced by the presence of a stem form. The regularity effect is not a result of the contexts, because it does not occur in the present tense. Because no effect of

Figure 3.14: Mean RT's for regulars and irregulars for present, meaning, form, and simple conditions



regularity was found for the meaning condition, the results are inconsistent with an account where regular and irregular past tenses differ when produced straight from meaning. The greatest contributor to the RTs in the meaning condition is the effect of context and not the effect of regularity. There is a non-significant 19 ms regularity effect in the meaning condition, and this study cannot prove that regularity plays *no* role in the production of past tense verbs from meaning, but it does prove that stem exposure increases the effect of regularity.

3.6.1 Frequency X Regularity interaction

Frequency did not have the expected effects in any of the conditions. For the present, subliminal form, and form conditions, this could be explained by the primacy of constraining context over frequency. Frequency effects do not occur when the word to be produced is highly constrained by context. However, in the simple condition, there was no context given, and the effect of frequency was still not as expected. In previous studies using this task (Prasada, Pinker, and Snyder, 1990; Seidenberg and Bruck, 1990) the result of interest was that high frequency irregulars were faster than low frequency irregulars, while high and low frequency regulars did not differ as shown in Figure 3.15.

For Prasada, Pinker, and Snyder (1990) this was taken as evidence that irregulars were stored, and thus subject to frequency effects, while regulars were not stored, but created by rule, leaving them free of frequency effects. Seidenberg and Bruck (1990) interpreted the result differently, explaining the difference in terms of statistical properties of input-output pairs. This study did not replicate the effect. High frequency verbs were slower than low frequency verbs while high frequency regulars were faster than low frequency regulars, as shown in Figure 3.16.

Figure 3.15: Expected shape of Frequency X Regularity interaction

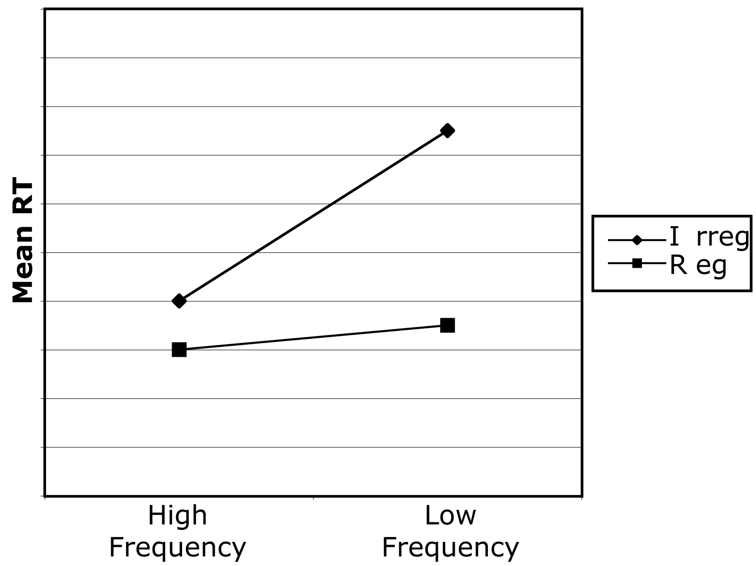
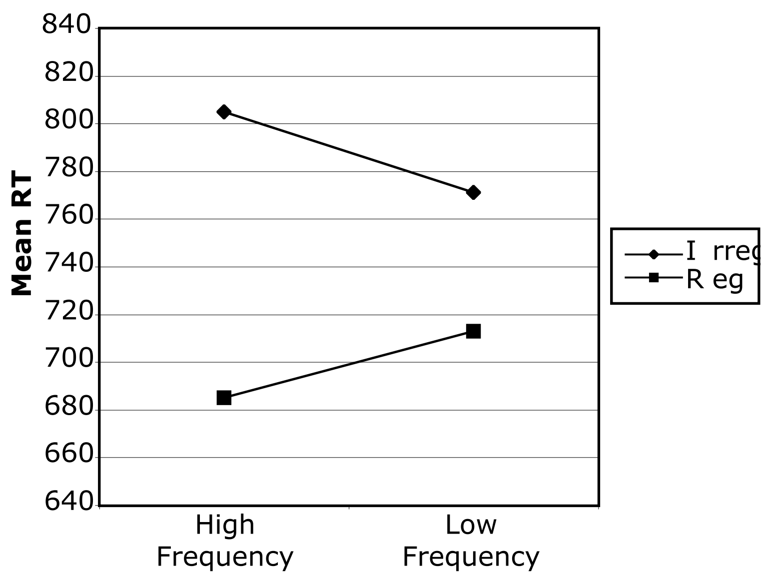


Figure 3.16: Frequency X Regularity interaction for simple condition



While this interaction was not significant in all the conditions, all the conditions did exhibit the same interaction shape as shown in Figure 3.16. This suggests that there may be some aspect of the words apart from regularity or frequency that contributed a bias in reaction times. The shape of the interaction was the same by subject and by item, so it was unlikely that particular items or subjects were biasing the results. Possibilities for bias in the phonetic properties of the words were considered. Because voice onset times may differ depending on the phonemes in the beginning of the word, target words were analyzed in terms of different properties of their initial phonemes including voicing, sibilancy, and presence of cluster. No biases were discovered. Because the regular past tense has an added syllable when the stem ends in a dental, and that could add production time, target words were analyzed in term of their final phonemes (dental or not) and no biases could be found. Stemberger (1993, 2004; Stemberger and Middleton, 2003) finds that something he calls 'vowel dominance' can affect access to past tense words. Vowel dominance has to do with the relationship between the stem vowel and the past tense vowel of a lexeme. The overall frequency of particular phonetic features throughout the lexicon will affect which vowel in an stem-irregular pair will more likely to be produced, so that in a regular-irregular pair like *sleep-slept*, the stem form will be dominant over the past form because it has a [+high] vowel. No biases of vowel dominance were discovered.

The selection of the target items was constrained by the meaningful context, which required a high level of intersubject agreement in the stimulus development phase, so frequency was not manipulated as carefully as it was in other studies, although frequencies were balanced between regular and irregular groups. Stem frequency was not controlled for, as it was in the Prasada, Pinker, and Snyder

(1990) study, by selecting verbs that had the same cluster frequencies, but different past tense frequencies. In this study high frequency verbs were higher in both past tense and cluster frequency than low frequency verbs, so it is still rather surprising that high frequency verbs should be slower.

It is possible that the low frequency verbs chosen were not low frequency enough. Some of the very low frequency irregulars, like *sought* were too difficult to elicit with the meaning task (what context predicts the target *sought* but not *looked for*?). It could be that the difference in frequency was not great enough. It could also be that the Francis and Kučera (1983) frequency measures were not the right ones to use. The verbs were recategorized in terms of the American Heritage word frequency index (Carroll, Davies, and Richman, 1971). This moved one verb from high frequency, to low frequency (*swung*) and 12 verbs from low frequency to high frequency (*sold, ate, sang, bit, taught, cut, filled, washed, rolled, pushed, stretched, crossed*). Under this analysis, low frequency verbs did take longer than high frequency verbs in the other conditions, but not in the simple condition, where it is the most expected because of the lack of context influence.

Because the two papers that originally found the frequency by regularity interaction were never-published conference papers, there was no comprehensive explanation of methods and stimuli I could examine in order to compare my stimuli to them and see where they differed. Because it cannot be evaluated, it is possible that the effect does not really exist. This is unlikely because it was also found in an error analysis of natural speech data (Stemberger and MacWhinney, 1988). For some reason my stimuli do not obtain the expected result with respect to frequency. I have not been able to figure out why, but since the stimuli were specifically designed to work with the meaning task, and not specifically designed to explore the

Table 3.7: Number of errors by type for each condition

	Over regularization	No marking	False irreg	Past/pres	Voice key	Wrong lexeme	Total # of trials
Present	0	1	0	22	24	141	1368
Meaning	3	0	0	N/A	18	190	1440
SubForm	1	1	0	N/A	16	100	1296
Form	1	2	0	N/A	34	0	1080
Simple	13	1	9	N/A	40	0	1080

frequency by regularity interaction, they do not say one way or the other whether that interaction exists. My hunch is that my low frequency words were not low frequency enough, so my contrast between high and low frequency was not as great as that of the other studies.

3.6.2 Error analysis

In the experiments six types of errors were made. Three types related specifically to regularity, and the other three did not. The regularity errors were those categorized as overregularization, no marking, or false irregular. They relate to regularity in that they are produced as intended irregular forms, but they are not the correct irregular forms. The other three error types, past instead of present, voice key, and wrong lexeme, do not relate to regularity in any particular way. The number of errors of each type for each condition is shown in Table 3.7.

A response was coded as an overregularization error if the subject produced a regular past tense (*digged*) in place of the correct irregular past tense (*dug*). The response was coded as a no marking error if they responded with the stem (*spend*) rather than a past tense (*spent*). False irregular errors only occurred in the simple task of Experiment 4. In these cases the subject responded with a nonce form (*shew*) in place of the correct past tense form (*showed*), or they responded with an actual

Table 3.8: Rates for errors related to regularity for each condition

	Meaning	SubForm	Form	Simple
Regularity errors	3	2	3	23
Total valid responses	1304	1180	1046	1040
Error rate	.002	.002	.003	.02

English word (*thank*) instead of the correct past tense form (*thought*). Past instead of present errors only occurred in the present tense condition of Experiment 2. In these cases, the subject said the past tense when they were supposed to say the present tense. A response was coded as a voice key error if the subject's response did not trip the voice key. It was coded as a lexeme error if instead of the target word (*walked*) they produced a context-appropriate word that was not the intended target (*went*). These errors were made because the contexts did not predict their targets 100%. No wrong lexeme errors were made in the form and simple conditions because the lexeme was explicitly presented in its stem form.

Error rates are shown in Table 3.8. Error rates were computed as the number of errors related to regularity divided by the total number of valid responses. Valid responses are the total number of responses with errors not related to regularity subtracted out.

The 2% error rate in the simple stem to past transformation task was almost ten times the error rate in any of the other tasks. This error rate was significantly different from those of all of the other conditions by Fisher's exact test, $p < .0001$, two-tailed. The error rates in the other three past tense conditions are too low to tell whether there are any differences between those conditions.

Researchers interested in investigating phenomena by using error rates as measurements often use controlled experiments that increase the likelihood of errors.

One reason people make more errors in controlled experiments than they do in natural situations is that in experiments there is often time pressure. When people are required to respond quickly, they make mistakes. Another way experiments encourage errors is to present stimuli that may interfere with processing as in cognitive load or Stroop paradigms. People might also make more errors in experiments because they do not have support from natural context cues that may provide helpful information in the real world. For example, in natural situations, people usually read the word *broil* within a recipe, and so understand it readily, but isolated in an experiment, *broil* may be more difficult to understand. Errors related to verb regularity do happen in natural speech (Stemberger and MacWhinney, 1988) but they are much more likely to occur in experimental situations for the reasons above.

Why did the simple condition in Experiment 4 lead to so many more errors? The first reason, time pressure, may have led to some errors in this experiment. However, time pressure would not fully explain the higher error rate in the simple condition because time pressure was also present in the other conditions. The stimulus sentence would disappear if the subject didn't respond quickly enough. The time pressure was slightly higher in the simple condition because the response was required earlier in the trial (only one word was shown instead of three sentences), but subjects always had a 6 second window in which to respond.

The second reason, interference stimuli, may have had some effect on error rates. The interference stimulus in this case would be the verb stem. The stem form may conflict with the past form the subject is preparing to say. However, this still doesn't explain the high error rate for the simple condition because the stem was also interfering with the past in the form condition and few errors were made. Additionally, one might expect that interference from the stem form would lead to

Table 3.9: False irregulars produced in the simple condition

Stem	False irregular past
show	shew
fill	fell
fix	fought
shave	shouve
drive	driv
think	thank
set	sat
feel	fell
run	rung

a large number of overregularization errors—the subject has been primed by the stem, and is therefore more likely to respond with the stem, simply adding the $-ed$ ending. While 57% of the errors made in the simple condition were indeed of this type, there were also 30% errors of the false irregular type. The false irregular errors are shown in Table 3.9. These types of errors cannot be explained simply by stem priming. Instead of repeating the stem form, the subject creates a form based on analogy with other past tenses.

The third reason, lack of supporting context, does set the simple condition apart from the other conditions. It was the only condition where context sentences were not given. Because context could not drive the choice of the past tense, subjects were explicitly told to produce the past tense of the words they saw. The combination of the lack of context and the explicit instruction to transform the tense together create a situation where the metalinguistic aspects of the task are highlighted. When people produce words in context, it is the context that drives the choice of words; they may or may not be aware of the tense properties of those words. Such a task is more linguistic than metalinguistic because they are not forced to be aware of the linguistic properties of the word in order to say the word. In the simple task, because there is no context, their chief motivation for producing a word is the awareness of its tense properties. This is a metalinguistic task. Subjects are aware of being in 'past tense' mode. This may be the reason for the increased number of errors. The inclination is to produce the stem with */ed/* on the end, because this is the most common thing to do, but this leads to overregularizaion errors. They know that the irregular past tenses can trip them up with this strategy. They compensate by being hypervigilant for the irregular pasts, resulting in false irregularization of words that don't have irregular pasts.

A closer look at the false irregular errors emphasizes the importance of meaning in past tense production. Some of the false irregulars produced are actual words. These words share form properties with the stimuli that elicited them, but no meaning properties. *Think-thank, fix-fought, fall-fell, feel-fell, run-rung*, these pairs all have phonemes in common, but their semantics are quite different. The lack of meaningful context in this condition leads to errors that seem to ignore meaning entirely. In the *think-thank* case, the word produced does not even have the semantic property of past tense. It is interesting that with respect to reaction time the simple condition is the same as the form condition, but with respect to errors, they are different. The tasks are the same in that they include a presentation of the stem form; they differ in that one presents meaning while the other does not. These false irregulars are rare in natural speech production (Stemberger, personal communication). In natural speech there is usually some meaning involved in the production of utterances. It appears that the input of meaning is a crucial part of normal past tense production. The dual and single route models presented in Chapter 1 leave meaning out of the input altogether, focusing solely on form input as illustrated in Figure 2.2. The results here lend support to models that incorporate semantics in the production process, like that of Joanisse and Seidenberg (1999). Production straight from form, with no semantic input, results in odd behavior, as reflected in the error rates of the simple condition.

3.7 Summary and general discussion

The four behavioral experiments discussed in this chapter, provide evidence that generating the past tense from form is different from generating the past tense from meaning. In the following section I summarize the findings from the four experiments and discuss the importance of these results for theoretical models of regularity in verb inflection.

The experiments together included five conditions that ranged from a pure meaning to form mapping task, to a pure form to form mapping task. The same regular and irregular verbs were produced in each task. No difference was found between regulars and irregulars in the present task of Experiment 2, where subjects produced the verbs in the present tense from present tense contexts. No difference was expected, because in the present tense all verbs are regular. However, this condition confirmed that there were no biases in the contexts used to elicit the verbs. Any regularity effects found would not be a result of variability in how easily the target words followed from their contexts.

No difference was found between regulars and irregulars in the meaning task of Experiment 1 either, which was the same as the present task, except that past tense verbs were produced from past tense contexts. Because past tense verbs are not all regular, one might have expected an effect of regularity in the condition, but none was found. A marginal difference was found in the subliminal form task, also of Experiment 1, where a masked prime of the stem was presented. This suggested that exposure to the stem form creates a regularity effect. Because conscious awareness of the prime may have been a factor in creating this effect, Experiment 3 used the form task, where the stem was presented as a fully visible prime. There was a significant difference between regulars and irregulars found in this condition.

Finally, in Experiment 4, a straight form to form mapping task also resulted in a significant difference between regulars and irregulars. This was the task used by previous researchers (Prasada, Pinker, and Snyder, 1990; Jaeger et al., 1996; Seidenberg and Bruck, 1990) as support for claims about the nature of regular and irregular inflection. The present results call into question the appropriateness of using this task to make such claims, since it appears the relevant differences between regulars and irregulars are caused by particular qualities of the task, namely the presentation of the stem form (but also perhaps the metalinguistic nature of the task), and not by underlying language processing mechanisms.

The interaction between frequency and regularity found in previous studies (Prasada, Pinker, and Snyder, 1990; Seidenberg and Bruck, 1990; Stemberger and MacWhinney, 1988; Ellis and Schmidt, 1998) was not replicated in this study. It is not clear why. The frequency characteristics of the stimuli could not be carefully controlled due to the nature of the meaning task. The primary concern in the development of the stimuli was that the target words be highly predictable from their elicitation contexts. The stimuli chosen happened to be equally distributed between high and low frequency, but it may have been the case that the low frequency words were not low frequency enough as a group.

Finally an error analysis of the conditions sets the simple condition apart from the others. Both the high error rates and types of errors in this condition implicate the importance of meaning in past tense production. Form to form transformation, without the input of meaning, is a feature of the dual and single route models discussed in Section 2.1. The form condition, which included meaning as input, resulted in fewer errors for past tense production than the simple condition, which excluded meaning as input. This suggests that models of past tense production should not leave out the input of semantics.

This study shows that generating a past tense from meaning is not the same as generating it from a stem form. The dual and single-route models of inflection as they stand do not account for this difference. The single-route model of Joanisse and Seidenberg (1999), because it incorporates a representation of meaning in the inflection process, has the potential to account for this difference, but so far it has only been used to model behavior in the form to form task. The dual-route model does not give a role to semantics in inflection, and does not explain why the behavior it aims to capture only occurs under conditions where the stem form has been presented. The results of this study require that models of verb inflection account not only for regularity effects, but also for the way those effects are changed by exposure to a stem form.

Chapter 4

MRI experiment

4.1 Introduction

Past behavioral studies of regular and irregular inflection found differences in the response time to produce regular and irregular verbs. This result has generally been taken as evidence for the dual-route theory of inflection. However, those studies presented some form of the verb stem in order to elicit the past tense. The study presented in Chapter 3 compared conditions where a stem form was presented to conditions where a stem form was not presented and discovered that the difference between regulars and irregulars is affected by the presence or absence of the stem form. Only when the stem form is presented do regulars and irregulars appear to behave as two functionally distinct processes.

Functionally distinct cognitive processes are often implemented in distinct neural systems, and indeed this idea motivated Jaeger et al.'s PET study of verb inflection (1996, page 457). They found that regulars and irregulars had distinct activation patterns in the brain and took this finding as evidence that regular and irregular inflection are distinct cognitive processes—support for the dual-route model. Subsequent imaging studies also found distinct activation patterns and interpreted them as evidence for a dual-route model. But Seidenberg and Hoeffner (1998) challenge the assumption that separate patterns of activation entail separate cognitive mechanisms. A single mechanism may rely on subsystems, like phonological processing, semantic processing, memory processing, etc., to differing *degrees* in different situations. A researcher may discover separate neural activation patterns if the degree of reliance on any one of those subsystems differs significantly between the situations tested.

However, whether a finding of two brain activation patterns is explained by one mechanism or two, all of the studies that produced this finding used a stem form to elicit the inflected form. It may be the case that regulars and irregulars only behave as neurally distinct processes when the stem form has been presented.

In this chapter I describe an experiment using functional magnetic resonance imaging (fMRI) to record brain activity during the production of past tense verbs using the same sentence completion stimuli used in the behavioral experiments of Chapter 3. Those behavioral experiments found that generating the past tense from a stem form is different from generating the past tense directly from meaning with respect to response time. This experiment is designed to test the hypothesis that generating the past tense from a stem is different from generating it from meaning with respect to brain activity. If so, previous imaging studies could be taken only as informative about stem-past mapping and not about inflection per se.

There were four behavioral experiments described in Chapter 3. Five different sets of subjects were run in five different conditions. In this fMRI experiment, only 2 conditions were used, the meaning condition, in which subjects see two context sentences and then a fill-in-the-blank sentence to which they respond with the appropriate past tense verb, and the form condition, which is the same as the meaning condition, but the stem form of the target verb is displayed in the blank. In this experiment, one set of subjects was tested in both conditions. This reduces general variability between subjects and, importantly, variability between individuals' brain anatomy. However, because subjects could only be exposed to a particular test item once, the number of trials for each condition was cut in half. This was considered the optimal way to transfer the behavioral experiment to an imaging experiment, because even if subjects had been able to do all 72 test items for each condition, the length of time they would have to spend in the magnet would be over the accepted limit for the subjects' comfort. Therefore the loss of some subject data was balanced by a reduction in scanner time and an elimination of between subject variability across conditions.

Adding more conditions to the experiment would have reduced even further the amount of data it was possible to collect, so only the two most important conditions were run. The meaning condition and the form conditions are the most important because they address the central question of this thesis: do generating a past tense from meaning and generating a past tense from form differ in the way they affect regulars and irregulars?

4.2 Methods

4.2.1 Subjects

Subjects were nine native speakers of English, 3 female and 6 male. All were students at the University of Chicago. All were strongly right handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). The subjects ranged in age from 23 to 34 and the mean age was 28.2. Subjects were paid for their participation.

4.2.2 Materials

The stimuli developed for the behavioral experiments of Chapter 3 were also used in this experiment. Each stimulus consisted of two context sentences followed by a fill-in-the-blank sentence. Of the stimuli used in the behavioral experiments, 33 were irregular and 36 were regular. (There had originally been 36 irregular verbs, but three were removed from analysis after it was discovered that they were outliers in terms of reaction time.) For the present experiment the regular and irregular verbs were each divided into two groups. The stem group would be presented with the verb stem in the blank, as in the form condition of Chapter 3, and the meaning group would be presented with an empty blank, as in the meaning condition of Chapter 3. Because there was an odd number of irregular verbs, one stimulus was added so that the stem and meaning groups would have equal numbers. The added item was for the irregular target *grew*: *His hair was quite short. Now it's at his shoulders because, it ____.* The added item had been part of the original list of 124 items given as a written test, and had scored high on intersubject agreement (95%), but had been eliminated from the experiment because it could not be easily used to elicit the present tense. Because the present tense would not be used in this experiment, it was added to the stimulus list.

Because there were fewer stimuli in each condition, it was important that errors be kept to a minimum so that all subject responses could be analyzed. Items that had higher rates of wrong lexeme errors in the behavioral experiment were put into the stem group. Because the stem would be presented for these items, there would be no indeterminacy about the target lexeme, and subjects would be less likely to make wrong lexeme errors. The categorization of regular and irregular verbs into stem and meaning groups is shown in Table 4.1. Because the comparison of interest

Table 4.1: Stimulus grouping for imaging experiment

Meaning Group		Form Group	
Irreg	Reg	Irreg	Reg
bite	ask	catch	cross
blow	brush	drink	fix
cut	change	drive	hire
dig	cry	eat	lie
fall	fill	grow	open
feed	fire	hear	paint
feel	help	hide	play
find	kiss	hit	push
forget	knock	run	roll
leave	lick	sell	stir
put	park	sing	stop
ring	point	slide	stretch
set	save	sweep	touch
sit	shave	swim	try
spend	show	swing	use
swear	sign	teach	walk
throw	wait	think	wash
	work		watch

in this experiment would be the difference between regulars and irregulars in the stem condition vs. the difference between regulars and irregulars in the meaning condition, it was confirmed that in both the stem group and the meaning group regulars and irregulars were balanced in terms of frequency, and number of syllables in second context sentence. They were also balanced in terms of concreteness and imageability of the target word (Coltheart, 1981). Additionally the present tense response times of behavioral Experiment 2 were used as a measure of how difficult it was to come up with a target word from its context. It was confirmed that in both groups regulars and irregulars were balanced with respect to accessibility of target word from its context.

Figure 4.1: Procedure for imaging experiment

	Meaning	Form
Fixation top of screen 2000 ms	+	+
Clause 1 2000 ms	There was one more place at the conference table.	There was one more place at the conference table.
Clause 2 2000 ms	He went to the chair.	He went to the chair.
Clause 3 2000 ms	and he ____ down.	and he <u>SAT</u> down.
Fixation center of screen 12000ms	+	+

4.2.3 Task protocol

Subjects saw a total of 70 stimuli, presented in random order in an event-related (single-trial) design. They saw 18 regular verb contexts and 17 irregular verb contexts in the meaning group, and 18 regular verb contexts and 17 irregular verb contexts in the form group. Stimuli were presented using E-Prime (2000) software, projected to a mirror above the subject's eyes in the scanner. The screen was divided into three frames of equal size. Each trial began with the presentation of a fixation point in the top of the top frame for 2 seconds, followed by the first context sentence presented for 2 seconds in the top frame, followed by the second context sentence presented for 2 seconds in the middle frame, followed by the fill-in-the-blank sentence for 2 seconds in the bottom frame. This was followed by a 12 second rest period where a fixation point was displayed in the middle of the screen. A schematic of the time course for each condition is shown in Figure 4.1.

Before the scan, subjects were given a practice session of 5 items. They were instructed to speak the word that most naturally fit the blank out loud when presented with the third sentence. They were told that sometimes a word would be in the blank, but it wouldn't necessarily be in the correct form for the context, and in that case they should say that word in the correct form. No mention was made of past tense. They were told that head movement could make their scan unreadable, and that speaking out loud often causes unwitting head movement. The experimenter then had them practice speaking while holding their heads still.

In the scanner, subjects wore a headset microphone so that responses could be recorded to digital audio tape.

4.2.4 Imaging protocol

Scans were performed at the Brain Research Imaging Center of the University of Chicago on a 3 Tesla scanner (GE Medical Systems, Milwaukee, WI) with a standard GE head coil. 30 5mm anatomical T1-weighted images were acquired in the sagittal plane for slice localization. During task performance, 30 5mm functional (T2*-weighted) images were acquired in the sagittal plane using BOLD contrast with one-shot spiral technique (Noll et al., 1995) using a matrix size of 64 X 64, and in-plane resolution of 3.75 X 3.75 mm. A TR of 2 seconds was used to collect a total of 1424 whole brain images in two runs of equal length. A high resolution 3d T1-weighted MPRAGE spiral volume scan was also collected.

4.2.5 Image analysis

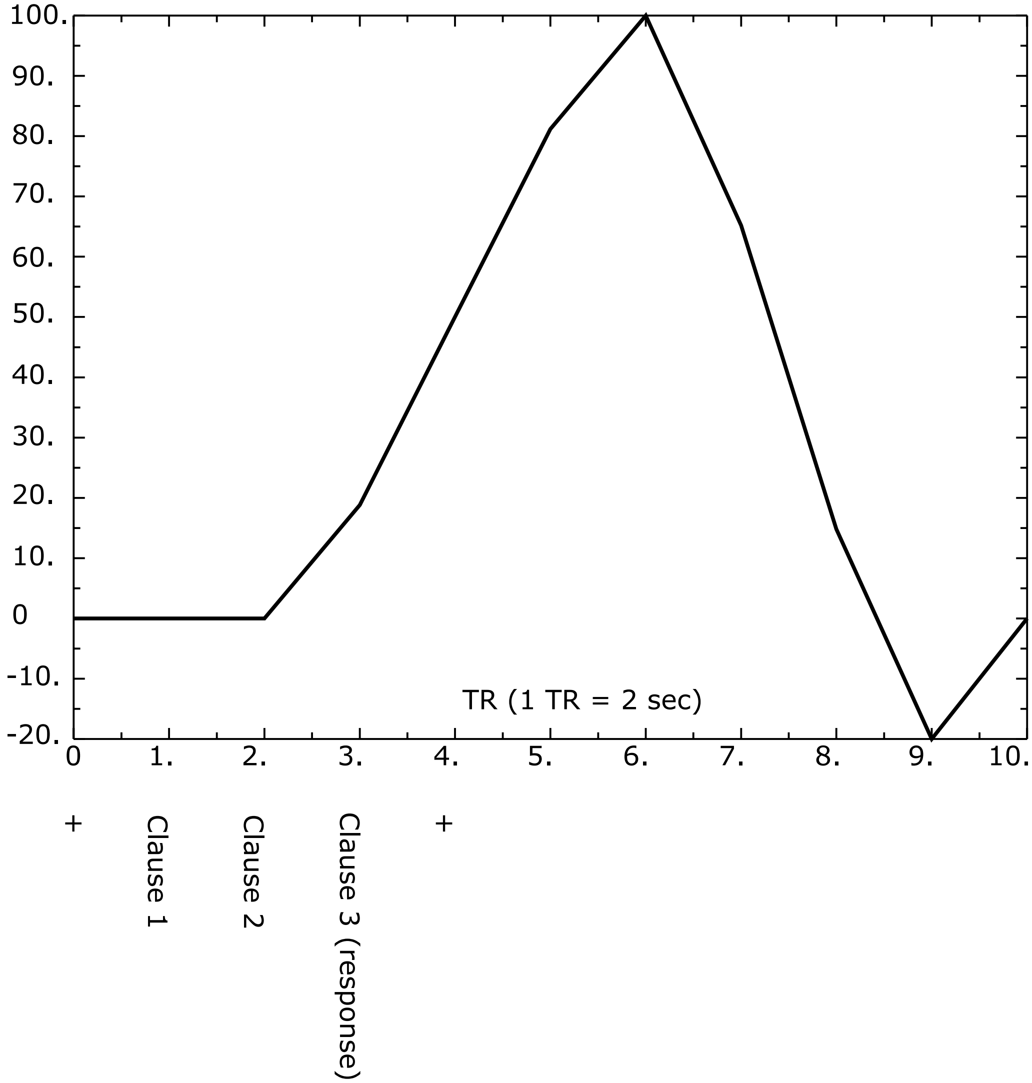
Images were spatially registered to check and correct for motion artifacts. No subjects moved more than 3 mm in any plane. Linear and quadratic trends were removed from the baseline signal of functional image series (Cox, 1996).

This experiment used an event-related design (Buckner et al., 1996; Josephs, Turner, and Friston, 1997; Zarahn, Aguirre, and D'Esposito, 1997). In block designs, the subject performs each experimental condition over a span of time that includes many trials of the same type, and brain activation over the entire span of time is compared between condition blocks. As Seidenberg and Hoeffner (1998) noted, this design was a problem for the Jaeger et al. (1996) study. In regular blocks the subjects repeated the same operation on every trial, but in the irregular blocks the subject had to produce a different type of irregular on each trial; regularity was likely confounded with difficulty and strategy effects. In an event-related design, brain activation is measured separately for each event of interest to the researcher, and the trials for different conditions can be presented in random order.

The event for this event-related design is somewhat complex. The event of interest for the experiment is the subject's production of the past tense word; however, when the subject produces the word, activation is not starting from baseline in language areas which have already been made active by the reading of the two context sentences. For this reason the event is modeled not just as the production of the verb, but as the entire trial: seeing initial fixation point, reading two context sentences, reading third sentence and saying verb. The peak brain activation of the event is modeled as occurring 6 seconds after the TR when the subject generates the verb. However, the rise to that peak starts does not start at baseline when the verb is generated, but when the initial fixation point is shown. The hemodynamic response to the stimuli was modeled as a 20 second waveform with a rise to peak value over 8 eight seconds (after a 4 second delay) followed by a 6 second fall to an undershoot of 20 percent of the peak, followed by a 2 second rise back to baseline. This waveform is basically the *Cox special* (Cox, 1996) hemodynamic response model, but with a longer delay and rise time. This model assumes the peak activation for each event occurs 6 seconds after the subject sees the fill-in-the-blank sentence, in accordance with the general time course of the BOLD (blood oxygenation level dependent) signal being measured. The response model is shown in Figure 4.2. The AFNI program 3dDeconvolve (Cox, 1996) was used to generate IRF files (estimated impulse response functions for each voxel) and in every subject IRF waveforms for speech-motor areas in the left precentral gyrus (assumed to be involved in producing the word) displayed the general shape of the response model described above.

Modeling the event in this way has the added benefit of reducing the amount of speech related artifact. Speaking can result in motion artifacts in the images collected by fMRI. For this reason language researchers use silent word generation tasks, as Beretta et al. (2003), and Ullman, Bergida, and O'Craven (1997) did in their past tense verb studies (Jaeger et al. and Indefrey et al. used PET, which is less susceptible to head movement). However, when the subject is asked to produce language silently, it is not possible to evaluate task performance. For this reason, overt language production is preferable if possible. Studies of overt speech in fMRI have found that signal changes associated with speech-related motion artifacts peak much earlier than the BOLD response and quickly rise from and return to baseline (Huang, Carr, and Cao, 2001; Birn et al., 1999). Because the hemodynamic response model used in this experiment rises slowly to a late peak, it will not correlate well with movement related artifacts, and therefore will not appear in thresholded activation maps. Also, because movement artifacts tend to occur in areas which lie on tissue boundaries and in lower slices near the throat and mouth (Barch et al., 1999), the ventricles and the cerebellum were left out of analysis.

Figure 4.2: Hemodynamic response model. The event occurring at each TR is shown along the x-axis. The peak of the model occurs 6 seconds after the subject responds.



Waveforms were created for each level of the conditions: irregular meaning, regular meaning, irregular stem, regular stem. Trials where the subject response was an error, or where the subject did not respond within the same TR as the fill-in-the-blank sentence were not used in analysis. The waveforms were used as regressors in a linear regression analysis of the functional time series (Ward, 2001). Activated voxels for each condition were determined using a single voxel statistical threshold of $F = 10.83$ ($p < .001$) and a cluster size of three or more contiguous voxels. Four activation maps were created, one for each level of the conditions.

In order to see which areas were activated by this task overall a general map was created, that showed which voxels were active for all subjects in all conditions.

For between subjects analysis, activation maps were transformed into Talairach space (Talairach and Tournoux, 1993) and resampled at a voxel size of $3mm^3$, and clusters were blurred by convolution with a Gaussian function of 9 mm (FWHM). Two voxel-wise paired t tests yielded two maps: one that showed areas of activation where regulars and irregulars differed in the meaning condition, and one that showed areas of activation where regulars and irregulars differed in the form condition.

Additionally, a region of interest analysis was carried out in which the brain was divided into 24 regions, 12 in each hemisphere. The specific regions are named in Tables 4.2 – 4.8. The mean intensity of activation in each region was determined for each subject's thresholded activation map, and those means were entered into a paired t test comparing irregulars to regulars in both the meaning and the form conditions.

4.3 Results

4.3.1 Behavioral results

The average number of errors was 3 (4% of responses) and no subject had more than 6 errors (8%). Three types of error were made: wrong lexeme, no response, and late response. No past tense errors were made (i.e., overregularization). All but 3 of the total errors were made in the meaning condition, but these were divided equally between regular and irregular verbs.

4.3.2 Imaging results

In general, this task activated areas areas known to be involved in reading and in the production of speech (see Turkeltaub, 2002, for overview). As shown in Appendix B.1, all conditions activated ventral occipital-temporal areas, more extensively on the left, and including the lingual gyrii and fusiform gyrii, areas known to be involved in visual aspects of reading. There was also activation in the posterior superior temporal gyrii, slightly more on the left, the precentral gyrii and the supplementary motor area, areas known to be involved in speech production. Surprisingly, activation in the inferior frontal gyrus, known to be crucial for language processing in the left hemisphere (e.g., Binder et al., 1997), was found primarily in the right hemisphere for this task. This may have to do with the more complex discourse processing required in this task, as opposed to the standard language tasks that use single words as stimuli. Right hemisphere brain damage typically leaves the ability to access words and sentences in tact, but impairs processes that require integration over a larger time scale like discourse based inferences (Brownell and Martino, 1998). Beeman, Bowden, and Gernsbacher (2000) found that the right

Table 4.2: Areas active for irregulars in the form condition (voxel-wise t test) as shown in Figure B.2

Coordinates	Hemisphere	Location (extension into other locations)
<u>Frontal</u>		
14 32 38	Right	Medial Frontal gyrus (Mid Frontal and Superior gyrii, cingulate and caudate)
-16 34 32	Left	Superior Frontal gyrus (Medial and Middle Frontal gyrii)
-10 34 44	Left	Superior Frontal gyrus (Medial Frontal gyrus)

hemisphere had greater involvement specifically in the processing of *predictive* inferences in connected discourse. The task used in this experiment requires subjects to make predictive inferences based on the context sentences. They must integrate meaning over the first two stimulus sentences in order to predict the most likely word for the blank in the third sentence, and this may lead to greater involvement of the right hemisphere language homologues.

The voxel-wise t tests revealed areas where regulars and irregulars differed. Voxels that met a significance value of $p < .01$ (uncorrected for multiple comparisons) were considered activated, but, in order to reduce the possibility of Type I error, only if they belonged to clusters of adjacent voxels totaling 20 or more. The areas where regulars and irregulars differed were not the same for the form and the meaning tasks. Areas of activation are shown in Appendices B.2 and B.3 and in Tables 4.2 and 4.3.

For the form task, irregular verb responses produced more activation than regulars in anterior clusters covering portions of the superior, medial and middle frontal gyrii in both hemispheres. The activation cluster in the right hemisphere also extended into the cingulate/caudate region. For the meaning task, there were no areas

Table 4.3: Areas active for regulars in the meaning condition (voxel-wise t test) as shown in Figure B.3

Coordinates	Hemisphere	Location (extension into other locations)
Frontal		
-38 16 30	Left	Middle Frontal gyrus (Precentral gyrus)
-34 28 6	Left	Inferior Frontal gyrus (Middle Frontal)
-14 2 48	Left	Cingulate (Medial Frontal)
22 44 20	Right	Superior Frontal gyrus (Middle and Medial Frontal)
Motor (and adjacent)		
52 -2 42	Right	Precentral gyrus (Middle Frontal)
-22 -28 56	Left	Postcentral gyrus (Precentral gyrus)
Parietal/Occipital		
22 -68 44	Right	Superior Parietal Lobule (Cuneus)
-14 -80 30	Left	Cuneus (Precuneus)
-14 -64 6	Left	Lingual gyrus (Cuneus)
46 -46 -6	Right	Fusiform gyrus (Middle Temporal gyrus)
32 -86 20	Right	Middle Occipital gyrus (Cuneus)

where irregular responses produced more activity than regular responses. Regulars produced more activation than irregulars in many areas: clusters in the left middle and inferior frontal gyrus, the right superior frontal gyrus, the left cingulate, the right and left precentral and left postcentral gyrii, the right superior parietal lobule, the left cuneus and lingual gyrus, and the right fusiform and middle occipital gyrii.

The voxel-wise t tests were performed on clusters of activation that had been resampled to a larger voxel size and blurred. Blurring is done in analyses of group data because brain anatomy between subjects varies considerably, and although the brains have been translated into a common coordinate system, activated voxels in the same anatomical regions in multiple subjects may not overlap in the group space. Upsampling and blurring increases the likelihood that they will overlap and produce significant results. However, the blurred activation may spread into anatomical regions other than the one in which it originated, and yield false positives for that region.

Because of this, a more conservative region of interest analysis was also carried out. Each subject's brain, in talairach space, was divided into 24 regions of interest, 12 in each hemisphere. Using the $1mm^3$ unblurred individual activation maps for each subject for each condition, the mean intensity of activation for each region of interest was computed. These means were then entered into t tests comparing irregulars to regulars in the form condition, and irregulars to regulars in the meaning condition. The results are presented in Table 4.4 and 4.5.

As in the voxel-wise t test, irregular responses generated more areas of significant activation than regular responses in the form condition, but there were no areas where regular responses produced more activation. The frontal area was significantly more active for the irregulars bilaterally, and the inferior frontal gyrus

Table 4.4: P values for mean activation by ROI for irregulars in the form condition.

	Left	Right
Frontal		
Middle, medial and superior frontal gyrus	$p < .05$	$p < .05$
Inferior frontal gyrus, 44 and 45		$p < .05$
Anterior cingulate		
Motor		
Precentral gyrus		
Postcentral gyrus		
Temporal		
Superior temporal gyrus		
Middle and inferior temporal gyrus		
Parietal		
Superior and inferior parietal lobule		
Occipital		
Cuneus, precuneus and occipital gyrii		
Fusiform gyrus		
Lingual gyrus		

Table 4.5: P values for mean activation by ROI for regulars in the meaning condition. Marginally significant results are marked with an asterisk.

	Left	Right
Frontal		
Middle, medial and superior frontal gyrus		$p < .05$
Inferior frontal gyrus, 44 and 45	$p < .05$	
Anterior cingulate	$p < .05$	
Motor		
Precentral gyrus	$p < .05$	$p < .05$
Postcentral gyrus		$p < .01$
Temporal		
Superior temporal gyrus	$p < .07$*	
Middle and inferior temporal gyrus		$p < .01$
Parietal		
Superior and inferior parietal lobule		$p < .05$
Occipital		
Cuneus, precuneus and occipital gyrii	$p < .05$	$p < .05$
Fusiform gyrus		$p < .01$
Lingual gyrus	$p < .05$	

was significantly more active on the left. The bilateral frontal activation was also found in the voxel-wise t test, but the inferior frontal activation had not been. This may be because the cluster size threshold used in the t test was too high to catch that activation, or because the voxels within the inferior frontal region in the ROI analysis that were active were not close enough to each other to overlap in the t test.

In the meaning condition, in contrast to the form condition, regulars produced more activation than irregulars, and there were no areas where irregulars were more active than regulars. In the right hemisphere the frontal region, the pre and post central gyrii, the lower temporal gyrii, the parietal area, the cuneal/occipital area, and the fusiform gyrus had greater activation for the regulars. In the left hemisphere the inferior frontal gyrus, anterior cingulate, precentral gyrus, cuneal/occipital areas, and lingual gyrus were more active. The superior temporal gyrus was more active to a marginally significant degree. Most of the activation in this analysis matched what was found in the voxel-wise t test except for 2 areas. The superior temporal gyrus was active for the ROI analysis (though to marginal significance) but not for the t test. The left middle frontal gyrus was active in the t test but not in the ROI analysis. That middle frontal gyrus activity may have originated in the precentral gyrus (which was active in the ROI analysis) but ended up with a center of mass in an adjacent area due to blurring.

In order to insure that it was not the case that *any* comparison of stimuli from this experiment would yield results, regardless of regularity or stem exposure, the stimuli were recategorized randomly into four groups, each containing the same number of stimuli as the original conditions. The data were then analyzed based on this recategorization. Two paired t tests were run where two of the new categories

Table 4.6: Stimuli eliminated from second analysis

Elimnated items	
Irreg	feel
	put
	spend
Reg	help
	kiss
	park
	wait

were compared to each other. Neither test revealed any ROIs with significant differences in activation between groups. Nor were any results marginally significant (the lowest p value for any area was $p = .15$).

These results are striking in how differently regulars and irregulars pattern in the meaning and the form conditions. The greater frontal activation for irregulars in the form condition replicates previous results using form to form mapping tasks. However, the greater activation for regulars in the meaning condition raises questions, especially, why should regulars activate occipital visual areas more than irregulars? This particular result implicated a possible visual difference between the stimuli.

The stimuli were checked for differences in visual characteristics, and while they had been balanced for the length of the second clause and the fill-in-the-blank sentence, it was discovered that the regulars in the meaning condition did have an average of 1.5 more words in the first clause than the irregulars did. This was an unfortunate oversight, but reanalysis with a balanced set of stimuli was still possible. The removal of seven stimuli from the meaning condition (three from irregular, four from regular) left the remaining test items balanced on number of words. The analysis was rerun with the remaining test items. The removed items are indicated in Table 4.6.

Results are shown in Appendix B.4 and Tables 4.7 and 4.8. The reanalysis of

Table 4.7: Areas active for regulars in the corrected meaning condition (voxel-wise t test) as shown in Figure B.4

Coordinates	Hemisphere	Location (extension into other locations)
Frontal		
22 44 20	Right	Superior Frontal gyrus (Medial Frontal)
Motor		
50 -8 38	Right	Precentral gyrus
-34 -10 38	Left	Precentral gyrus
Temporal		
44 -34 2	Right	Superior Temporal gyrus (Middle Temporal)
Occipital		
-16 -76 38	Left	Precuneus (Cuneus)
-16 -64 6	Left	Lingual gyrus (Cuneus)
44 -40 -12	Right	Fusiform gyrus (Inferior Temporal gyrus)
-40 -58 -4	Left	Middle Occipital gyrus (Fusiform)

Table 4.8: P values for mean activation by ROI for regulars in the meaning condition, after balancing for clause 1. Marginally significant results are marked with an asterisk.

	Left	Right
Frontal		
Middle, medial and superior frontal gyrus		.07*
Inferior frontal gyrus, 44 and 45		
Anterior cingulate		
Motor		
Precentral gyrus	$p < .07$*	$p < .05$
Postcentral gyrus		$p < .05$
Temporal		
Superior temporal gyrus	$p < .07$*	
Middle and inferior temporal gyrus		$p < .07$*
Parietal		
Superior and inferior parietal lobule		
Occipital		
Cuneus, precuneus and occipital gyrii	$p < .05$	
Fusiform gyrus		$p < .05$
Lingual gyrus	$p < .05$	

the meaning condition compares only 14 regular to 14 irregular trials. This is sparse data for an imaging study, but certain results remain. Regulars still had more areas of activation than the irregulars, and there were no areas where irregulars had more, as was the case in the original analysis. Active areas were similar. Some areas that had shown regulars with more activation in the original analysis did not show as more active in the reanalysis: the left inferior frontal gyrus and anterior cingulate, the right parietal lobe, and the right cuneus/occipital area. Some areas only reached marginal significance in the reanalysis: the right frontal area, left precentral gyrus, and the right middle/inferior temporal area. Again the t test and ROI analyses were well matched. The left superior temporal gyrus, marginally active in the ROI analysis, was not active in the t test and the right superior temporal gyrus, active in the t test was not active in the ROI analysis, though that cluster also extended into the middle temporal gyrus which was marginally active for ROI. Again, blurring may have shifted the center of mass to an adjacent area.

4.4 Discussion

In the stem condition, the irregular responses produced more activity in the frontal regions, as was also found by Indefrey et al. (1997), Ullman, Bergida, and O'Craven (1997), Jaeger et al. (1996), and Beretta et al. (2003). Ullman, Bergida, and O'Craven attribute this activity to lexical search for irregulars, Jaeger et al. attribute it to response inhibition of the regular rule, and Beretta et al. attribute it to working memory processes for lexical search of irregulars (Indefrey et al. do not give a hypothesis). However, irregular responses did not activate this area more than regulars in the meaning condition, suggesting that the activation is not specific to regularity. What else may account for this greater frontal activity? One aspect that all the other studies shared with the stem condition of this study is that another form of the target word was presented as part of the stimulus. The greater frontal activation for irregulars in the form condition seems to be related to the production of irregulars only when the subject has viewed the stem. When the subjects see the stem and must say the irregular past, they must inhibit the phonological form of the stem in order to say the irregular correctly. When the stem is not shown, this inhibition is not necessary. This account is consistent with the behavioral results of Chapter 3. Because the the regular responses do not require the inhibition of the stem form being presented, they are produced faster than the irregulars. The greater frontal activation may be due to response inhibition for saying the stem itself (which would lead to overregularization), as Jaeger et al. suggested. It less clear how lexical search or working memory, as Ullman and Beretta suggested, would be involved more only when the stem has been presented (it seems that the lack of a stem cue rather than the presence of one would place more demands on lexical search and working memory), but whatever the source of the activation difference, because it was not found in the meaning condition, it appears not to be specific to the irregularity of the verbs.

It is also possible that the difference in activation for the form condition is a result of facilitation for the regulars due to priming. The most consistent finding of brain imaging studies of priming is known as "repetition suppression" (see Henson, 2003; Schacter and Buckner, 1998, for review). Primed stimuli yield a reduced neural response compared to unprimed stimuli. This reduction in response "reflects faster or 'more efficient' processing of the primed stimulus, owing to performance of the same processes in the recent past (on the prime)" (Henson, 2003, page 57). Repetition suppression has been found specifically in left prefrontal regions when the stimuli are processed at a conceptual (as opposed to surface perceptual) level (Buckner et al., 1997; Demb et al., 1995; Wagner et al., 1997; Wagner et al., 2000). Seeing the stem may have primed the production of regulars better than irregulars.

Beretta et al. (2003) suggested the priming explanation, but only in the sense that an abstract regular inflection operation may have primed subsequent performance of this operation. They reject a word level interpretation of a priming effect because each specific word was only shown one time. However, it could be the case that a specific word primes not just a repeated presentation of that word, but also a production of that word. Having seen the word, the subject requires fewer neural resources to generate the word (or a phonologically similar inflected form of the word).

In the meaning condition, regulars were more active than irregulars, contrary to all other studies. Why would regulars have more activation? Ullman's theory predicts that they would involve more procedural and rule processing areas in the inferior frontal gyrus and basal ganglia (Ullman, 2001). Neither of these areas was found to be more active for regulars. However regulars did activate the precentral gyrus bilaterally and the left posterior superior temporal gyrus (to marginal significance). The voxel-wise t test showed the precentral gyrus activation to be in the middle portion of the gyrus commonly activated by movement of the mouth lips and tongue in fMRI studies (Fox et al., 2001; Cao et al., 1993; Huang, Carr, and Cao, 2001). This portion of the precentral gyrus controls the motor articulation of speech, and the posterior superior temporal gyrus has been implicated in phonemic encoding of words for articulation (Hickok, 2001; Hickok et al., 2000) and has been shown to vary in activation as a function of phonological load in articulation (Okada et al., 2003). The motor and posterior superior temporal gyrus activity taken together are consistent with an interpretation where the regulars are more active because they are more difficult to articulate. Some single route proponents have claimed that neuropsychological dissociations between regular and irregular verbs can be attributed to the phonological complexity of regulars. The addition of the /d/t/ phoneme for regular past tenses means they are usually more phonologically complex, and therefore more difficult to articulate, than irregulars, (Burzio, 2002; Bird et al., 2003; McClelland and Patterson, 2002).

Interestingly, there are still visual areas more active for the regulars even though the stimuli are now balanced on amount of reading necessary. It is not obvious how to interpret this visual activation. It is possible that a visual imagery strategy was used to a greater degree for the production of regulars in the meaning task. Visual imagery is known to activate occipital areas (Kosslyn et al., 1999). In the meaning task, the subject was looking at a blank where the past tense verb should be. They may have been imagining the visual form of the word in the blank. If so, they were imagining longer words for the regulars than for the irregulars. While the verb stems of the regulars and irregulars contain the same average number of letters, the past tenses for the regulars are an average of two letters longer than the irregular past tenses because of the addition of the $-ed$ ending. The active visualization of longer words in the regular meaning condition may have led to greater activation in visual areas.

4.5 Summary and general discussion

The imaging experiment discussed in this chapter presents preliminary evidence that the brain response to generating regular and irregular past tense verbs differs depending on whether they are generated from form or from meaning. When the stem form of the verb is presented to the subject, irregulars activate more areas than regulars, particularly in the frontal gyrii. This result is consistent with previous imaging studies, all of which included the verb stem as part of the stimulus. When the subject must generate the verb solely from meaning, regulars activate more areas than irregulars, particularly in motor and visual areas. This is inconsistent with previous findings, and suggests that those findings were dependent on the task characteristic of stem exposure.

The experiment presented here suffered from a few methodological problems. In order to adapt the experiment to the requirements of fMRI protocols, the number of trials presented was quite low. The discovery of an imbalance in the number of words presented in clause 1 of the protocol necessitated the further reduction of the number of trials. In averaged images for the voxel-wise t tests, a strict p

value, corrected for multiple comparisons, did not reveal any activation, so a lower, uncorrected p value was used to explore the data though combined with a high cluster size. A more strict ROI analysis did reveal areas of activation that differed between conditions. The discussion of these results offers explanations as to why the particular areas of difference were found, but further study will be necessary to explore those explanations as direct hypotheses.

The task used in this experiment required complex linguistic processing. Each trial involved form processing in both phonology and orthography, and meaning processing at the level of word, sentence and discourse. In addition demands were made on syntactic processing, memory, attention, and motor and visual systems. Imaging studies of language function usually try to specifically target only one type of processing in order to determine the location of that processing in the brain. Stimuli are typically single syllables or words (e.g., Zatorre et al., 1996), or single phrases or sentences (e.g., Bottini et al., 1994). The language used in such tasks is far from the natural use of language we encounter in daily life, but the circumscribed nature of those tasks allows researchers to posit specific mechanisms for specific functions. This experiment used a much less circumscribed task, but the purpose of the experiment was not to posit a specific mechanism for a specific function. Previous experiments had already found that regulars and irregulars appeared to be generated by separate mechanisms. The purpose was of this experiment was to determine whether a particular feature of those studies, stem exposure, itself altered the nature of difference between regulars and irregulars. The results suggest that it does. The mechanisms involved in verb inflection are affected by the presence or absence of a stem form.

What are those mechanisms and how are they altered? In order to answer this question, more targeted experiments are necessary. However, the pattern of results suggests some possible answers. The greater frontal activation for irregulars over regulars in the form task suggests that inhibition mechanisms are more important for irregulars when the verb stem is presented or possibly that priming effects are greater for regulars. The greater precentral activation for regulars in the meaning condition suggests that motor pronunciation mechanisms are more important for regulars. The greater visual activation suggests that visual imagery strategies are used to a greater degree for regulars when no stem is shown. The results are not straightforward to interpret, but they do offer one general important finding—the difference between regulars and irregulars in the form condition look nothing like that difference in the meaning condition.

Chapter 5

Conclusion

At the beginning of this thesis I referred to suggestions from Marcus et al. (1992), Marchman (1988), and Patterson et al. (2001) that a recent presentation of the stem form of a verb may affect production of regular and irregular past tenses differently. This study, in a direct comparison of past tense production with stem exposure to past tense production without stem exposure, shows that this is indeed the case. In verb inflection, form to form mapping is not the same as meaning to form mapping.

Chapter 2 introduced the competing theories of regularity and verb inflection, and described the behavioral and brain imaging studies that have been used to support them. Problems with form to form mapping were discussed, both in terms of the implicit assumptions of the theoretical models and the experimental tasks used to investigate them.

Chapter 3 first discussed the development of an experimentally controlled meaning to form mapping task. This task was used to elicit regular and irregular past tense verb forms. The response times to produce the verbs were compared to those from other tasks in which the stem of the verb was presented. Results showed that exposure to the stem form of the verb significantly increased the response time difference between regular and irregular verbs. No significant difference between regulars and irregulars was found in the meaning to form mapping task.

In Chapter 4, the behavioral experiment of the previous chapter was adapted for use in an fMRI study of brain activation in the production of past tense verbs.

Like previous imaging studies, it found that the production of irregular past tense verbs produced greater activation in frontal areas than regular past tense verbs, but only in the case where a stem form had been presented. In the meaning to form mapping task, irregulars did not produce more activation than regulars in any regions. Regulars produced more activation than irregulars in precentral and occipital areas, a result no previous study has obtained.

In future work, I hope to test the validity of these findings in other domains. One important extension of this study would be to use a language other than English. German, for example, has a class of irregular verb participles that have the same degree of phonological overlap with their stems as regular verbs. Similar findings in a study of German verb inflection using these irregulars would indicate that it is not just phonological overlap that enhances regular verb inflection or inhibits irregular verb inflection. Hebrew, a language where both regular and irregular inflection involve changing the vowel patterns interspersed with the consonants, also offers a test of the role of phonology in the difference between meaning to form and form to form mapping.

Another extension of this work would be to other populations. The effect of stem exposure was originally noticed in children and in brain damaged subjects. This effect was on error rates. In this study error rates in the form to form task where no context was given were greater and of a different type than in the tasks where context was given. It remains to be seen whether a lack of meaningful context affects children or brain damaged subjects in a similar way.

This thesis offers valuable contributions. First, it presents an experimental task for investigating the generation of verbs from meaning. Meaning can be an ambiguous topic, and it is notoriously difficult to control experimentally. The meaning to form task used here is more naturalistic than tasks that use simple stem forms as stimuli, while at the same time retaining experimental control. It provides a means of investigating verb inflection while avoiding any biases introduced by a pre-given phonological state.

Second, the imaging experiment overcomes some of the shortcomings of previous regularity experiments by using a event-related design rather than a block design, and by successfully using overt spoken responses.

Third, the general result of both the behavioral and imaging experiments, that exposure to the verb stem has an effect on the inflection process, is a new finding. Although this result itself is not evidence that supports or discounts either the single or dual-route model, it requires accommodation by any model that aims to account for human language use. A successful model should incorporate a representation of meaning, and be able to explain why a mapping from meaning to past tense differs from a mapping from stem to past tense.

Appendix A

Stimuli

A.1 Sentences and target words

Table A.1: List of stimuli

Clauses	Target	Stem	Frequency	Regularity
The frisbee was flying toward him.	caught	catch	H	Irreg
Upon putting out his hand,				
he ____ it.				
He was going to visit relatives in St. Louis.	drove	drive	H	Irreg
After getting in the car,				
he ____ there.				
He was trying to walk on a thin ledge.	fell	fall	H	Irreg
He couldn't keep his balance and				
he ____.				
The anesthetic wasn't working completely,	felt	feel	H	Irreg
because when they were putting the stitches in his hand,				
he ____ it.				

A.1, Stimuli, continued

Clauses	Target	Stem	Frequency	Regularity
He was looking for a stamp. He went through all the desk drawer, finally he ____ one.	found	find	H	Irreg
He was listening for the distant church bells. Finally, everyone went quiet and he ____ them.	heard	hear	H	Irreg
He was at bat. After missing the ball two times, finally He ____ it.	hit	hit	H	Irreg
The lecture was getting very boring. Nobody would care if he didn't stay. So, he ____.	left	leave	H	Irreg

A.1, Stimuli, continued

Clauses	Target	Stem	Frequency	Regularity
The wind was very cold. After taking his gloves out of his pocket, he ____ them on.	put	put	H	Irreg
He was only half a block from the bus stop. As it was starting to pull away, he ____.	ran	run	H	Irreg
There was one more place at the conference table. He went to the chair, and he ____ down.	sat	sit	H	Irreg
He was planning to get up with the alarm clock. Before he went to bed, he ____ it.	set	set	H	Irreg

A.1, Stimuli, continued

Clauses	Target	Stem	Frequency	Regularity
After finding an extra $30, he went to the mall, where he ____ it.	spent	spend	H	Irreg
He was trying to work on his follow-through. After picking up his golf club, he ____ it.	swung	swing	H	Irreg
He was away from home for a long time. His hometown was special to him. Often, he ____ about it.	thought	think	H	Irreg
He was playing fetch with his dog. After wrestling the stick from the dog's mouth, he ____ it.	threw	throw	H	Irreg

A.1, Stimuli, continued

Clauses	Target	Stem	Frequency	Regularity
He was very hungry. After taking a piece of bread from the bag, he ____ it.	ate	eat	L	Irreg
He was curious to see if it was real gold. Upon putting the ring between his teeth, he ____ it.	bit	bite	L	Irreg
He was chopping celery with a sharp knife. The knife went through his fingertip. His finger ____.	bled	bleed	L	Irreg
HIs birthday candles were dripping on the cake. The singing was finally over, so he ____ them out.	blew	blow	L	Irreg

A.1, Stimuli, continued

Clauses	Target	Stem	Frequency	Regularity
The wire was keeping the knob from turning. After getting the scissors, he ____ it.	cut	cut	L	Irreg
His throat was dry. After making some lemonade, he ____ it.	drank	drink	L	Irreg
He went to the spot where his coins were buried. After getting out his shovel, he ____ them up.	dug	dig	L	Irreg
His baby was making that hungry cry. After getting a bottle ready, he ____ her.	fed	feed	L	Irreg

A.1, Stimuli, continued

Clauses	Target	Stem	Frequency	Regularity
He was to bring his camera home from work. Both his wife and his son were reminding him all day. But still, he ____.	forgot	forget	L	Irreg
He was looking forward to skating. When it was cold enough to flood the yard with water, the water ____.	froze	freeze	L	Irreg
He didn't want his wife to find her birthday present. He went to the garage with it, where he ____ it.	hid	hide	L	Irreg
He was expecting an important call. He was sitting by the phone all day. Suddenly, it ____.	rang	ring	L	Irreg

A.1, Stimuli, continued

Clauses	Target	Stem	Frequency	Regularity
The audience was waiting to hear his famous baritone voice.	sang	sing	L	Irreg
As the music was starting, he went up to the microphone, and				
he ____.				
There was dog hair everywhere.	shed	shed	L	Irreg
Whenever it was spring again,				
the dog ____.				
The hardwood floor and his stockinged feet were giving him an idea.	slid	slide	L	Irreg
After taking a running start,				
he ____ across the room.				
He didn't need his old car anymore.	sold	sell	L	Irreg
After putting an ad in the paper,				
he ____ it.				

A.1, Stimuli, continued

Clauses	Target	Stem	Frequency	Regularity
He went to the health club to exercise. After easing into the pool, he ____.	swam	swim	L	Irreg
There was confetti all over the floor. After getting a broom, he ____ it up.	swept	sweep	L	Irreg
They didn't believe he was telling the truth. So after placing his hand on a bible, he ____ on it.	swore	swear	L	Irreg
His daughter was eager to learn to play tennis. They went to the courts every Saturday, and he ____ her.	taught	teach	L	Irreg

A.1, Stimuli, continued

Clauses	Target	Stem	Frequency	Regularity
He was wondering what college she went to.	asked	ask	H	Reg
So the next time they were talking,				
he ____ her.				
The elderly woman was having trouble taking her bags to her car.	helped	help	H	Reg
He was a gentleman, so				
he ____ her.				
There was a package waiting for him at home.	opened	open	H	Reg
After getting the tape off the lid,				
he ____ it.				
They were eager to hear his new composition.	played	play	H	Reg
So sitting down at the piano,				
he ____ it.				

A.1, Stimuli, continued

Clauses	Target	Stem	Frequency	Regularity
His son couldn't find Nepal on the map.	pointed	point	H	Reg
So he went to the map, and with his finger,				
he ____ to it.				
The kids were asking to see his tattoo.	showed	show	H	Reg
So, lifting the sleeve of his T-shirt,				
he ____ them.				
He was getting all green lights.	stopped	stop	H	Reg
But the next one was red, so				
he ____.				
He wasn't sure if those pants would fit him.	tried	try	H	Reg
so he went into the dressing room, and				
he ____ them on.				

A.1, Stimuli, continued

Clauses	Target	Stem	Frequency	Regularity
He was out of shampoo.	used	use	H	Reg
There was only a little bit in the bottle yesterday, and				
he ____ it up.				
He was at the dentist's office in time for his appointment,	waited	wait	H	Reg
but the dentist wasn't ready for him yet. So,				
he ____.				
He didn't live far from his office.	walked	walk	H	Reg
When the weather was nice,				
he ____ there.				
He was dying to see the big fight.	watched	watch	H	Reg
So he went to a bar with cable TV, and				
he ____ it.				

A.1, Stimuli, continued

Clauses	Target	Stem	Frequency	Regularity
There was a big report due on Monday. He went to the office on Saturday, and he ____ on it.	worked	work	H	Reg
He was concerned about keeping his teeth healthy. After lunch, and even after his coffee break, he ____ them.	brushed	brush	L	Reg
His pants were too casual to wear to dinner. He went back up to his room, and he ____.	changed	change	L	Reg
He didn't expect the experience to be so emotional. But when his daughter was born, he ____.	cried	cry	L	Reg

A.1, Stimuli, continued

Clauses	Target	Stem	Frequency	Regularity
His car was on the other side of the street. After looking both ways, he ____ it.	crossed	cross	L	Reg
There was only a little fuel left in his tank. He went to a gas station, and he ____ it up.	filled	fill	L	Reg
The company was asking everyone to cut their staff. After calling in the two least productive team members, he ____ them.	fired	fire	L	Reg
The garbage disposal wasn't working again. After getting all the necessary tools, he ____ it.	fixed	fix	L	Reg

A.1, Stimuli, continued

Clauses	Target	Stem	Frequency	Regularity
The fifth interview was the best applicant. After clearing it with his superiors, he ____ her.	hired	hire	L	Reg
He was so happy to be back in his native land. After dropping to his knees on the ground, he ____ it.	kissed	kiss	L	Reg
He wasn't sure if anyone was at home. He went up to the door, and he ____ on it.	knocked	knock	L	Reg
He was sealing envelopes for his party invitations. Instead of using a sponge to wet them, he ____ them.	licked	lick	L	Reg

A.1, Stimuli, continued

Clauses	Target	Stem	Frequency	Regularity
He didn't want his mother to know about his surgery.	lied	lie	L	Reg
When she was asking why he wasn't at work,				
he ____ to her.				
The view from the beach was breathtaking.	painted	paint	L	Reg
After getting his brushes and his easel,				
he ____ it.				
He was driving around the lot for a long time.	parked	park	L	Reg
Finally, there was an open spot, and				
he ____ there.				
His tires were spinning in the snow.	pushed	push	L	Reg
He went behind the car, and				
he ____.				

A.1, Stimuli, continued

Clauses	Target	Stem	Frequency	Regularity
They didn't put the brake on in the golf cart. The driveway was on a slight hill, and it ____ away.	rolled	roll	L	Reg
He was working to get enough for a house, so he didn't spend his paycheck, he ____ it.	saved	save	L	Reg
His wife was complaining about his scratchy face, so after getting a razor, he ____.	shaved	shave	L	Reg
He was in agreement with everything in the contract. After getting a pen from his pocket, he ____ it.	signed	sign	L	Reg

A.1, Stimuli, continued

Clauses	Target	Stem	Frequency	Regularity
The sugar was settling at the bottom of his teacup.	stirred	stir	L	Reg
So upon putting his spoon in the tea,				
he ____ it.				
He was getting ready for his workout.	stretched	stretch	L	Reg
To warm up, he went to the mats, and				
he ____.				
He didn't know the fence was electrified.	touched	touch	L	Reg
His hand was still smarting after, accidentally,				
he ____ it.				
His hands were dirty from working in the garden.	washed	wash	L	Reg
He went to the sink, and				
he ____ them.				

A.2 Frequency characteristics

FreqClass: Past tense forms are classified as high(H) or low(L) frequency based on a cutoff point of 35 per million in Francis and Kučera (1983). PastFreq: the absolute frequency of the past tense form. PresFreq: the absolute frequency of the the present tense form. ClustFreq: the frequency of the verb in all its forms. Clause2Syll: the number of syllables in the second context sentence for the verb. P values for t tests between the verb categories are reported. There are no significant differences between irregulars and regulars in any categories.

Table A.2: Characteristics of stimuli

IRREGULARS

Verb	Regularity	FreqClass	PastFreq	PresFreq	ClustFreq	Clause2Syll
caught	irreg	H	54	1	146	7
drove	irreg	H	58	5	203	7
fell	irreg	H	87	19	239	8
felt	irreg	H	302	45	643	13
found	irreg	H	268	58	1033	10
heard	irreg	H	129	7	433	10
hit	irreg	H	38	6	126	11
left	irreg	H	157	26	650	11
put	irreg	H	130	20	513	11
ran	irreg	H	134	16	431	9

A.2, Characteristics, continued

Verb	Regularity	FreqClass	PastFreq	PresFreq	ClustFreq	Clause2Syll
sat	irreg	H	139	6	314	6
set	irreg	H	71	14	372	6
spent	irreg	H	40	8	194	6
swung	irreg	H	43	0	77	8
thought	irreg	H	340	23	982	10
threw	irreg	H	46	5	150	10
ate	irreg	L	16	2	122	11
bit	irreg	L	7	2	26	10
bled	irreg	L	2	0	18	8
blew	irreg	L	12	5	52	10
cut	irreg	L	25	14	245	7
drank	irreg	L	19	3	93	8
dug	irreg	L	7	1	9	8
fed	irreg	L	8	7	132	9
forgot	irreg	L	17	0	119	15
froze	irreg	L	1	1	53	13
hid	irreg	L	6	1	61	9
rang	irreg	L	21	1	39	12
sang	irreg	L	28	9	120	16
shed	irreg	L	3	3	12	8
slid	irreg	L	24	0	43	8
sold	irreg	L	20	13	128	10
swam	irreg	L	6	0	55	8

A.2, Characteristics, continued

Verb	Regularity	FreqClass	PastFreq	PresFreq	ClustFreq	Clause2Syll
swept	irreg	L	19	0	54	6
swore	irreg	L	14	2	33	11
taught	irreg	L	19	11	153	11
Avg			64.17	9.28	224.25	9.47
AvgH			127.25	16.19	406.63	8.94
AvgL			13.70	3.75	78.35	9.90

REGULARS

Verb	Regularity	FreqClass	PastFreq	PresFreq	ClustFreq	Clause2Syll
asked	reg	H	300	17	612	8
helped	reg	H	40	27	352	7
opened	reg	H	94	16	259	9
played	reg	H	65	34	333	9
pointed	reg	H	48	19	143	11
showed	reg	H	138	72	640	9
stopped	reg	H	103	2	240	7
tried	reg	H	120	8	472	10
used	reg	H	137	32	1016	16
waited	reg	H	68	2	263	12

A.2, Characteristics, continued

Verb	Regularity	FreqClass	PastFreq	PresFreq	ClustFreq	Clause2Syll
walked	reg	H	143	7	287	6
watched	reg	H	68	1	209	12
worked	reg	H	76	34	496	11
brushed	reg	L	14	0	38	12
changed	reg	L	26	10	225	8
cried	reg	L	25	1	64	7
crossed	reg	L	26	3	84	6
filled	reg	L	31	5	184	8
fired	reg	L	19	0	78	14
fixed	reg	L	12	0	109	11
hired	reg	L	6	1	47	11
kissed	reg	L	15	1	31	10
knocked	reg	L	17	1	47	7
licked	reg	L	7	0	14	10
lied	reg	L	5	0	13	11
painted	reg	L	9	4	95	11
parked	reg	L	8	0	61	10
pushed	reg	L	31	2	102	7
rolled	reg	L	34	2	88	9
saved	reg	L	11	4	121	8
shaved	reg	L	4	0	23	8
signed	reg	L	15	2	62	10
stirred	reg	L	7	3	39	10

A.2, Characteristics, continued

Verb	Regularity	FreqClass	PastFreq	PresFreq	ClustFreq	Clause2Syll
stretched	reg	L	21	4	61	9
touched	reg	L	24	5	91	13
washed	reg	L	10	0	83	6
Avg			49.36	8.86	196.72	9.53
AvgH			107.69	20.85	409.38	9.77
AvgL			16.39	2.09	76.52	9.39
Ttest	Irreg vs. Reg		0.40	0.90	0.63	0.92
Ttest	IrregL vs. RegL		0.32	0.16	0.91	0.49
Ttest	IrregH vs. RegH		0.53	0.50	0.98	0.37

A.3 Frequency characteristics of revised list

The frequency characteristics of the stimuli target words after the removal of *bleed*, *freeze*, and *shed*.

Table A.3: Characteristics of revised stimuli

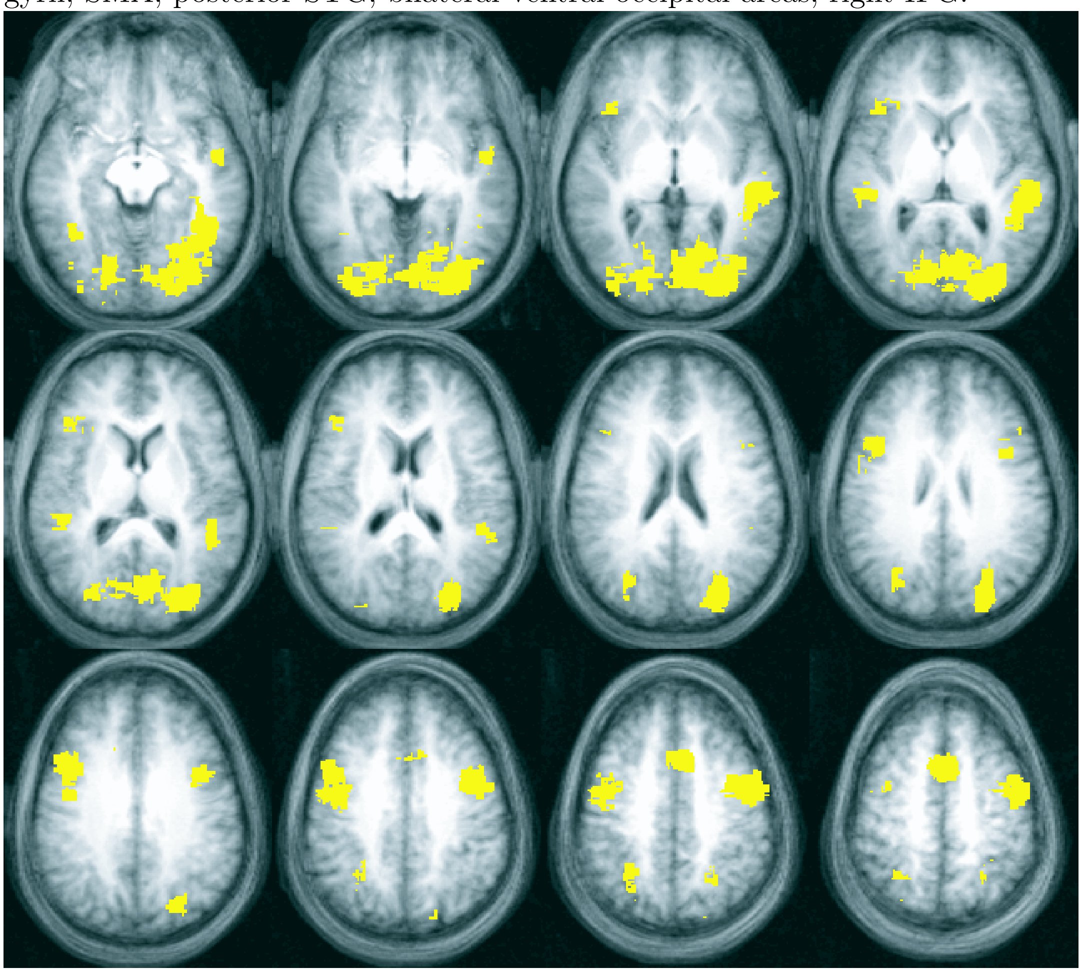
IRREGULARS		PastFreq	PresFreq	ClustFreq	Clause2Syll
Avg		69.82	10.00	242.12	9.45
AvgH		127.25	16.19	406.63	8.94
AvgL		15.76	4.18	87.29	9.94
REGULARS		PastFreq	PresFreq	ClustFreq	Clause2Syll
Avg		49.36	8.86	196.72	9.53
AvgH		107.69	20.85	409.38	9.77
AvgL		16.39	2.09	76.52	9.39
Ttest	Irreg vs Reg	0.27	0.73	0.44	0.90
Ttest	IrregL vs. RegL	0.81	0.11	0.55	0.49
Ttest	IrregH vs. RegH	0.53	0.50	0.98	0.37

Appendix B

Imaging pictures

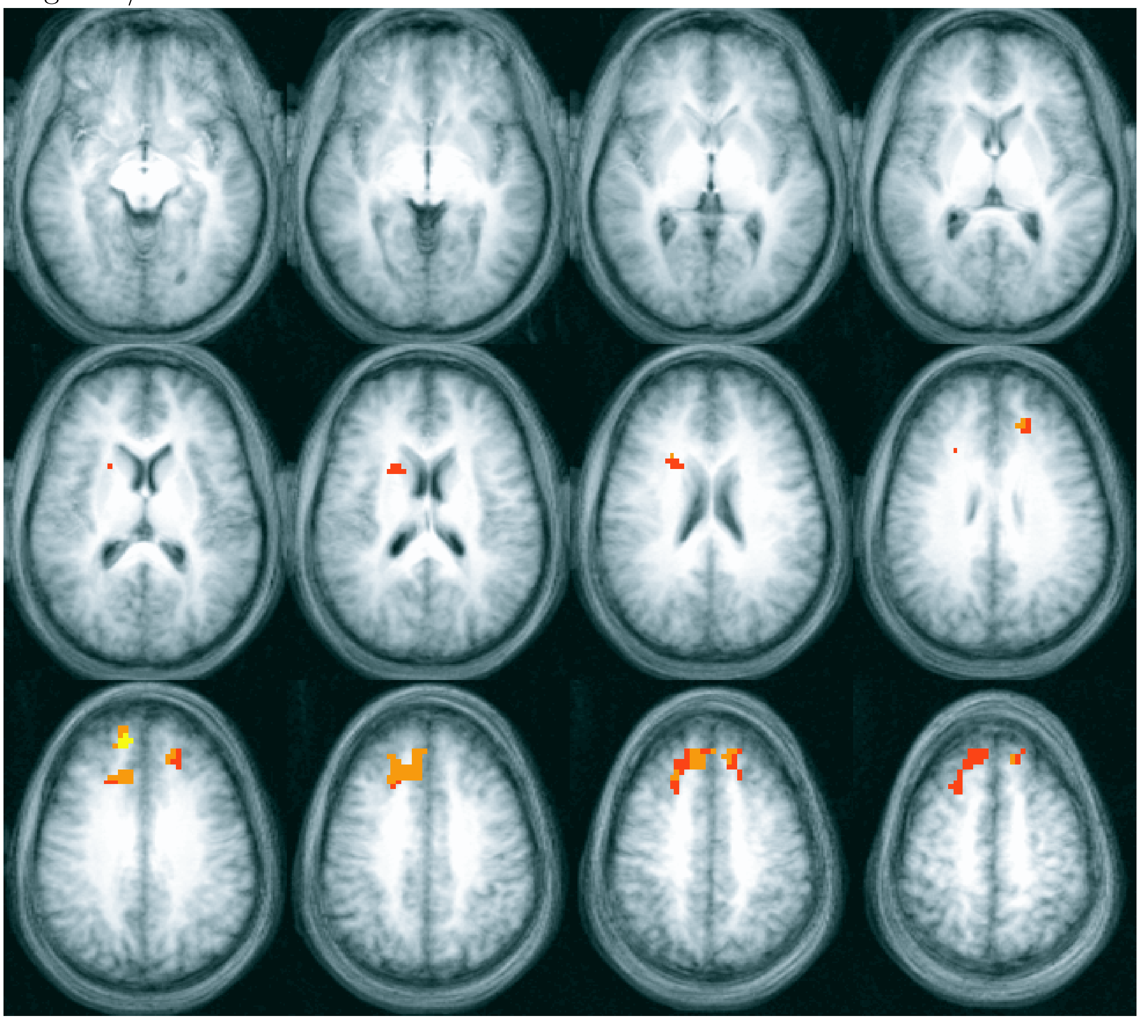
B.1 Voxels activated in all condtions

Figure B.1: Areas where subjects activated in all conditions. Bilateral precentral gyrii, SMA, posterior STG, bilateral ventral-occipital areas, right IFG.



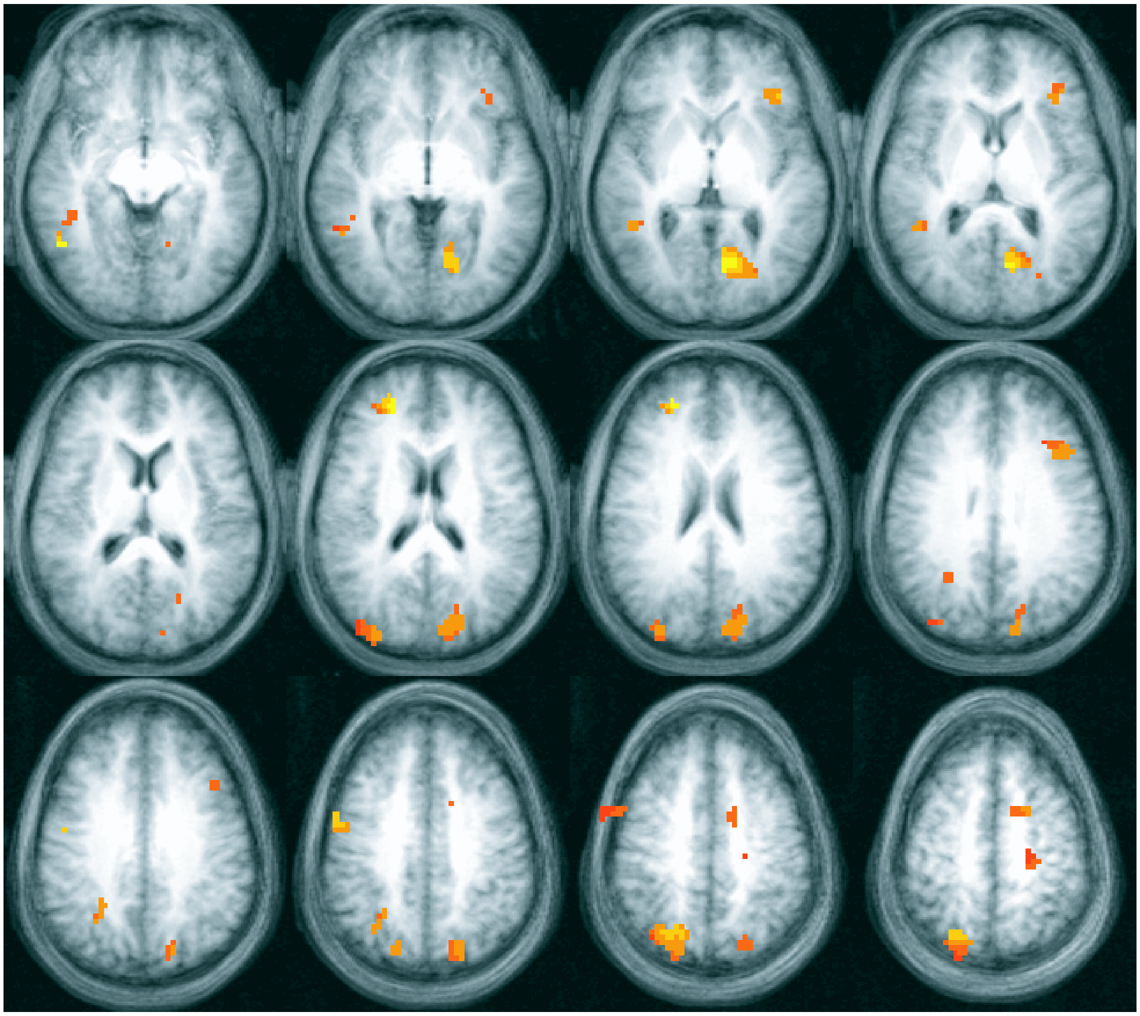
B.2 Form condition

Figure B.2: Results of t test for regulars vs. irregulars in form condition. Areas where irregulars were more active are in orange. There were no ares where regulars were more active. Bilateral superior, middle, medial frontal gyrus, right cingulate/caudate.



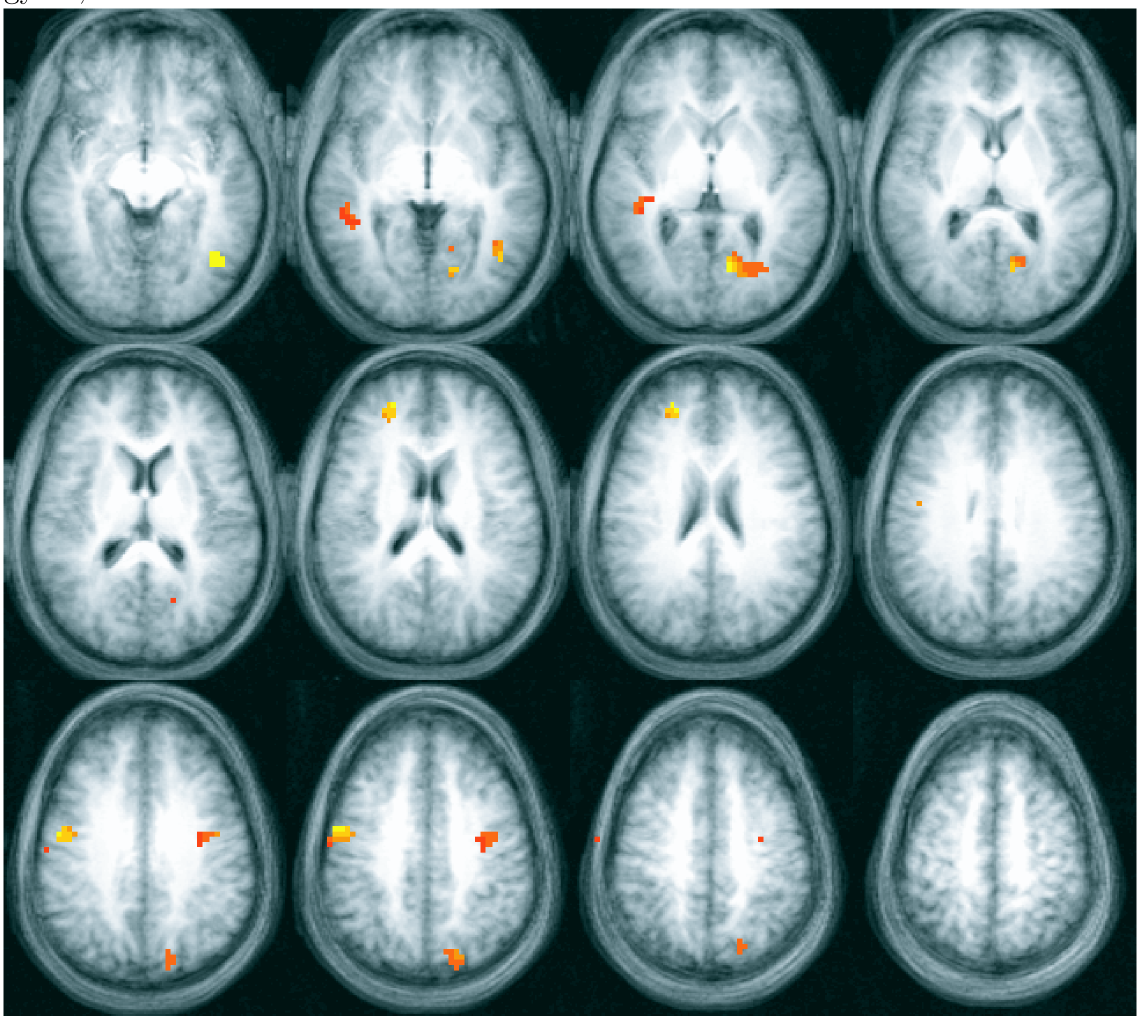
B.3 Meaning condition

Figure B.3: Results of t test for regulars vs. irregulars in the meaning condition. Areas where regulars were more active are in orange. There were no areas where irregulars were more active. Right SPL, right and left precentral, right and left cuneus and precuneus, left MFG, right SFG, left lingual gyrus, right fusiform gyrus.



B.4 Corrected meaning condition

Figure B.4: Results of t test for regulars vs. irregulars in the meaning condition after unbalanced stimuli were eliminated. Areas where regulars were more active are in orange. There were no areas where irregulars were more active. Left cuneus and precuneus, bilateral precentral gyrii, right SFG, left lingual gyrus, right fusiform gyrus, left MOG.



References

Alegre, M. and P. Gordon. 1999. Frequency effects and the representational status of regular inflections. *Journal of Memory and Language*, 40:41–61.

Anderson, J.R. 1982. Acquisition of cognitive skill. *Psychological Review*, 89:369–406.

Baayen, R., T. Dijkstra, and R. Schreuder. 1997. Singulars and plurals in dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, 37:94–117.

Barch, D., F. Sabb, C. Carter, T. Braver, D. Noll, and J. Cohen. 1999. Overt verbal responding during fmri scanning: Empirical investigation of problems and potential solutions. *NeuroImage*, 10:642–657.

Beeman, M.J., E.M. Bowden, and M.A. Gernsbacher. 2000. Right and left hemisphere cooperation for drawing predictive and coherence inferences during normal story comprehension. *Brain and Language*, 71:310–336.

Berent, I., S. Pinker, and J. Shimron. 1999. Default nominal inflection in hebrew: Evidence for mental variables. *Cognition*, 72:1–44.

Beretta, A., C. Campbell, T. Carr, J. Huang, L. Schmitt, K. Christianson, and Y. Cao. 2003. An er-fmri investigation of morphological inflection in german reveals that the brain makes a distinction between regular and irregular forms. *Brain and Language*, 85:67–92.

Berko, J. 1958. The child's learning of english morphology. *Word*, 14:150–177.

Bird, H., M.A. Lambon Ralph, M. Seidenberg, J. McClelland, and K. Patterson. 2003. Defecits in phonology and past tense morphology: What is the difference? *Journal of Memory and Language*, 48:502–526.

Birn, R., P. Bandettini, R. Cox, and R. Shaker. 1999. Event-related fmri of tasks involving brief motion. *Human Brain Mapping*, 7:106–114.

Brown, R. 1973. *A first language: the early stages*. Harvard University Press, Cambridge, Mass.

Brownell, H.H. and G. Martino. 1998. Deficits in inference and social cognition: The effects of right hemisphere brain damage on discourse. In M. Beeman and C. Chiarello, editors, *Right hemisphere language comprehension: Perspectives from cognitive neuroscience*. Erlbaum, Mahwah, NJ, pages 309–328.

Buckner, R.L., P.A. Bandettini, K.M. O'Craven, R.L. Savoy, S.E. Peterson, and M.E. Raichle et al. 1996. Detection of cortical activation during averaged single trials of a cognitive task using functional magnetic resonance imaging. *Proceedings of the National Academy of Sciences*, 93:14878–14883.

Buckner, R.L., W. Koustaal, D.L. Schacter, S.E. Petersen, M.E. Raichle, and B.R. Rosen. 1997. Fmri studies of item repetition during word generation. *Cognitive Neuroscience Annual Meeting*, 4:67.

Bunge, S., K. Ochsner, J. Desmond, G. Glover, and J. Gabrieli. 2001. Prefrontal regions involved in keeping information in and our of mind. *Brain*, 124:2074–2086.

Burzio, L. 2002. Missing players: Phonology and the past-tense debate. *Lingua*, 112:157–199.

Bybee, J.L. and D.I. Slobin. 1982. Development and use of the english past tense. *Language*, 58:265–285.

Cao, Y., V.L. Towle, D.N. Levin, and J.M. Balter. 1993. Functional mapping of human motor cortical activation by conventional mri at 1.5 t. *Journal of Magnetic Resonance Imaging*, 3:869–875.

Carroll, J.B., P. Davies, and B. Richman. 1971. *The American Heritage word frequeny book*. Houghton Mifflin, Boston, Mass.

Clahsen, H. 1999. Lexical entries and rules of language: A multidisciplinary study of german inflection. *Behavioral and Brain Sciences*, 22:991–1060.

Clahsen, H. and M. Rothweiler. 1992. Inflectional rules in children's grammars: Evidence from the development of participles in german. *Morphology Yearbook*, pages 1–34.

Clark, H.H. 1973. The language as fixed effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 12:335–359.

Cohen, G. and D. Faulkner. 1983. Word recognition: Age differences in contextual facilitation effects. *British Journal of Psychology*, 74:239–251.

Coltheart, M. 1981. The mrc psycholinguistic database. *Quarterly Journal of Experimental Psychology*, 33A:497–505.

Conway, A., S. Tuholski, R. Shisler, and R. Engle. 1999. The effect of memory load on negative priming: An individual differences investigation. *Memory and Cognition*, 27:1042–1050.

Cottrell, G. and K. Plunkett. 1995. Acquiring the mapping from meaning to sounds. *Connection Science*, 6:399–412.

Cox, R. 1996. Afni: Software for analysis and visualization of function magnetic resonance neuroimages. *Computers in Biomedical Research*, 29:162–173.

Cummings, J.L. 1993. Frontal-subcortical circuits and human behavior. *Archives of Neurology*, 50:873–880.

Daneman, M. and I. Green. 1986. Individual differences in comprehending and producing words in context. *Journal of Memory and Language*, 25:1–18.

Daugherty, K. and M. Seidenberg. 1992. The past tense revisited. In *Paper presented at The Proceedings of the 14th Annual Meeting of the Cognitive Science Society, Princeton, NJ.*

Davis, C., A. Castles, and E. Iakovidis. 1998. Masked homophone and pseudohomophone priming in children and adults. *Language and Cognitive Processes*, 13:625–651.

Dell, G.S., M.F. Schwarz, N. Martin, E.M. Saffran, and D.A. Gagnon. 1997. Lexical access in normal aphasic speakers. *Psychological Review*, 104:801–838.

Dellwo, V. and P. Wagner. 2003. Relations between language rhythm and speech rate. In *15th ICPhS, Barcelona*, pages 471–474.

Demb, J.B., J.E. Desmond andd A.D. Wagner, C.J. Vaidya, G.H. Glover, and J.D. Gabrieli. 1995. Semantic encoding and retrieval in the left inferior prefrontal corex: a functional mri study of task difficulty and process specificity. *Journal of Neuroscience*, 15:5870–5878.

2000. E-prime 1.0. Pittsburgh, PA. Psychology Software Tools. [computer software].

Ellis, N.C. and R. Schmidt. 1998. Rules or association in the acquisition of morphology? the frequency by regularity interaction in human and pdp learning of morphosyntax. *Language and Cognitive Processes*, 13:307–336.

Ferrand, L. and J. Grainger. 1992. Phonology and orthography in visual word recognition: Evidence from masked nonword priming. *Quarterly Journal of Experimental Psychology*, 45A:353–372.

Ferrand, L. and J. Grainger. 1994. Effects of orthography are independent of phonology in masked form priming. *Quarterly Journal of Experimental Psychology*, 47A:365–382.

Forster, K.I. 1999. The microgenesis of priming effects in lexical access. *Brain and Language*, 68:5–15.

Fowler, C.A., S.E. Napps, and L. Feldman. 1985. Relations among regular and irregular morphologically related words in the lexicon as revealed by repetition priming. *Memory and Cognition*, 13:241–255.

Fox, P., A. Huang, L.M. Parsons, J.-H. Xiong, F. Zamarippa, L. Rainey, and J.L. Lancaster. 2001. Location-probability profiles for the mouth region of human primary motor-sensory cortex: Model and validation. *NeuroImage*, 13:196–209.

Francis, N. and H. Kučera. 1983. *Frequency analysis of English usage: Lexicon and grammar*. Houghton Mifflin, Boston, Mass.

Garrett, M.F. 1975. The analysis of sentence production. In G.H. Bower, editor, *The psychology of learning and motivation*. Academic Press, New York, pages 133–177.

Gentner, D. 1981. Some interesting differences between nouns and verbs. *Cognition and Brain Theory*, 4:161–178.

Goldman-Eisler, F. 1958. The predictability of words in context and the length of pauses in speech. *Language and Speech*, 1:226–231.

Goldman-Rakic, P.S. 1987. Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In Fred Plum, editor, *Handbook of physiology: nervous system, volume 5: Higher functions of the brain (part 1)*. Bethesda, MD: American Physiological Society, pages 373–417.

Gonnerman, L.M. 1999. *Morphology and the lexicon: Exploring the semantics-phonology interface*. Ph.D. thesis, University of Southern California, Los Angeles, CA.

Griffin, Z.M. 2002. Recency effects for meaning and form in word selection. *Brain and Language*, 80:465–487.

Griffin, Z.M. and K. Bock. 1998. Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language*, 38:313–338.

Grosjean, F. and J. Itzler. 1984. Can semantic constraint reduce the role of word frequency during spoken word recognition? *Bulletin of the Psychonomic Society*, 22:180–182.

Hahn, U. and R. Nakisa. 2000. German inflection: Single route or dual route? *Cognitive Psychology*, 41:313–360.

Hare, M. and J. Elman. 1995. Learning and morphological change. *Cognition*, 56:61–98.

Hare, M., J. Elman, and K. Daugherty. 1995. Default generalization in connectionist networks. *Language and Cognitive Processes*, 10:601–630.

Henson, R.N.A. 2003. Neuroimaging studies of priming. *Progress in Neurobiology*, 70:53–81.

Hickok, G. 2001. Functional anatomy of speech perception and speech production: Psycholinguistic implications. *Journal of Psycholinguistic Research*, 30:225–235.

Hickok, G., P. Erhard, J. Kassubek, A.K. Helms-Tillery, S. Naeve-Velguth, J.P. Strupp, P.L. Strick, and K. Ugurbil. 2000. An fmri study of the role of left posterior temporal gyrus in speech production: Implications of the explanation of conduction aphasia. *Neuroscience Letters*, 287:156–160.

Huang, J., T. Carr, and Y. Cao. 2001. Comparing cortical activations for silent and overt speech using event-related fmri. *Human Brain Mapping*, 15:39–53.

Indefrey, P., C. Brown, P. Hagoort, H. Herzog, M. Sach, and R.J. Seitz. 1997. A pet study of cerebral activation patterns induced by verb inflection. *NeuroImage*, 5:S548.

Jaeger, J.J., A.H. Lockwood, D.L. Kemmerer, R.D. Van Valin, B.W. Murphy, and H.G. Khalak. 1996. A positron emission tomographic study of regular and irregular verb morphology in english. *Language*, 72:451–497.

Joanisse, M.F. and M.S. Seidenberg. 1999. Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Science USA*, 96:7592–7597.

Josephs, O., R. Turner, and K. Friston. 1997. Event-related fmri. *Human Brain Mapping*, 5:243–248.

Kempen, G. and P. Huijbers. 1983. The lexicalization process in sentence production and naming: Indirect election of words. *Cognition*, 14:185–209.

Kempley, S.T. and J. Morton. 1982. The effects of priming with regularly and irregularly related words in auditory word recognition. *British Journal of Psychology*, 73:441–454.

Kosslyn, S.M., A. Pascal-Leone, O. Felician, S. Camposano, J.P. Keenan, W.L. Thompson, G. Ganis, K.E. Sukel, and N.M. Alpert. 1999. The role of area 17 in visual imagery: Convergent evidence from pet and rtms. *Science*, 284:167–170.

Kouider, S., R. Peereman, and E. Dupoux. Phonological activation during reading: Time-course or conscious awareness? Manuscript submitted for publication.

Kuczaj, S.A. 1977. The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior*, 16:589–600.

Lachman, R. 1973. Uncertainty effects on time to access the internal lexicon. *Journal of experimental psychology*, 99:199–208.

Lavric, A., D. Pizzagalli, S. Forstmeier, and G. Rippon. 2001. A double-dissociation of english past-tense production revealed by event-related potentials and low-resolution electromagnetic tomography (loreta). *Clinical Neurophysiology*, 112:1833–1849.

Levelt, W.J.M. 1989. *Speaking: From intention to articulation.* MIT Press, Cambridge, MA.

Ling, C. and M. Marinov. 1993. Answering the connectionist challenge: A symbolic model of learning past tenses of english verbs. *Cognition*, 49:235–290.

Lukacs, A., M. Racsmany, and C. Pleh. 2001. Vocabulary and morphological patterns in hungarian children with williams syndrom: A preliminary report. *Acta Linguistica Hungarica*, 48:243–269.

Lukatela, G., S. Frost, and M. Turvey. 1998. Phonological priming by masked nonword primes in the lexical decision task. *Journal of Memory and Language*, 39:666–683.

MacKay, D.G. 1976. on the retrieval and lexical structure of verbs. *Journal of Verbal Learning and Verbal Behavior*, 15:169–182.

MacWhinney, B. and J. Leinbach. 1991. Implementations are not conceptualizations: Revising the verb learning model. *Cognition*, 40:121–157.

Maratsos, M. 2000. More overregularizations after all: New data and discussion on marcus, pinker, ullman, hollander rosen and xu. *Journal of Child Language*, 27:183–212.

Marchman, V. 1988. Rules and regularity in the english past tense. *Center for research in language newsletter*, 2(4):San Diego.

Marchman, V. 1997. Children's productivity in the enlgish past tense: The role of frequency, phonology, and neighborhood structure. *Cognitive Science*, 21:283–304.

Marchman, V., K. Plunkett, and J. Goodman. 1997. Overregularization in english plural and past tense inflectional morphology: a response to marcus (1995). *Journal of Child Language*, 24:767–779.

Marcus, G. 1995. Children's overregularization of english plurals: A quantitative analysis. *Journal of Child Language*, 22:447–460.

Marcus, G., U. Brinkmann, H. Clahsen, R. Wiese, A. Woest, and S. Pinker. 1995. German inflection: The exception that proves the rule. *Cognitive Psychology*, 29:189–256.

Marcus, G., M. Ullman, S. Pinker, M. Hollander, T.J. Rosen, and F. Xu. 1992. Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*, 57.

Marslen-Wilson, W., L. Tyler, R. Waksler, and L. Older. 1994. Morphology and meaning in the english mental lexicon. *Psychological Review*, 101:3–33.

Marslen-Wilson, W.D., M. Hare, and L. Older. 1993. Inflectional morphology and phonoogical regularity in the english mental lexicon. In *Proceedings of the 15th annual meeting of the Cognitive Science Society*. Lawrence Erlbaum Associates Inc., pages 693–698.

McClelland, J.L. and K. Patterson. 2002. Rules or connections in past-tense inflections: What does the evidence rule out? *TRENDS in Cognitive Sciences*, 6:465–472.

Miller, E. and J. Cohen. 2001. An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24:167–202.

Murrell, G.A. and J. Morton. 1974. Word recognition and morphemic structure. *Journal of Experimental Psychology*, 102:963–968.

Nakisa, R. and U. Hahn. 1996. Where defaults don't help: The case of the german plural system. *Paper presented at the 18th Annual Conference of the Cognitive Science Society, San Diego.*

Napps, S.E. 1989. Morphemic relationships in the lexicon: Are they distinct from semantic and formal relationships? *Memory and Cognition*, 17:729–739.

Newman, A., R. Izvorski, L. Davis, H. Neville, and M. Ullman. 1999. Distinct electrophysiological patterns in the processing of regular and irregular verbs. *Journal of Cognitive Neuroscience*, S47.

Noll, D., J. Cohen, C. Meyer, and W. Schneider. 1995. Spiral k-space mri of cortical activation. *Journal of Magnetic Resonance Imaging*, 5:49–56.

Okada, K., K. Smith, C. Humphries, and G. Hickok. 2003. Word length modulates neural activity in auditory cortex during covert object naming. *NeuroReport*, 14:2323–2326.

Oldfield, R.C. 1971. The assessment and analysis of handedness: The edinburgh inventory. *Neuropsychologia*, 9:97–113.

Orsolini, M., R. Fanari, and H. Bowles. 1998. Acquiring regular and irregular inflection in a language with verb classes. *Language and Cognitive Processes*, 13:425–464.

Oscar-Berman, M., P. McNamara, and M. Freedman. 1991. Delayed-response tasks: Parallels between experimental ablation studies and finding in patients with frontal lesions. In Harvey S. Levin, Howard M. Eisenberg, and Arthur L. Benton, editors, *Frontal lobe function and dysfunction*. New York: Oxford University Press, pages 230–255.

Patterson, K., M.R. Lambon, J. Hodges, and J. McClelland. 2001. Deficits in irregular past-tense verb morphology associated with degraded semantic knowledge. *Neuropsychologia*, 39:709–724.

Penke, M., H. Weyerts, M. Gross, E. Zander, T. Münte, and H. Clahsen. 1997. How the brain processes complex words: an event-related potential study of german verb inflections. *Cognitive Brain Research*, 6:37–52.

Pinker, S. 1984. *Language learnability and language development*. Harvard University Press, Cambridge, MA.

Pinker, S. 1991. Rules of language. *Science*, 253:530–535.

Pinker, S. 1999. *Words and rules: The ingredients of language*. Harper Collins, New York.

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

Pinker, S. and A. Prince. 1994. Regular and irregular morphology and the psychological status of rules of grammar. In S.D. Lima, R.L. Corrigan, and G.K. Iverson, editors, *The reality of linguistic rules*. Benjamins, Philadelphia.

Plunkett, K. and V. Marchman. 1991. U-shaped learning and frequency effects in a multi-layered perceptrom. *Cognition*, 39:43–102.

Plunkett, K. and V. Marchman. 1993. From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition*, 48:21–69.

Plunkett, K. and R. Nakisa. 1997. A connectionist model of the arabic plural system. *Langauge and Cognitive Processes*, 12:1–47.

Prasada, S., S. Pinker, and W. Snyder. 1990. Some evidence that irregular forms are regtrieved from memory but regular forms are rule generated. *Paper presented at the 31st annual meeting of the Psychonomic Society (November: New Orleans).*

Prasada, S. and S. Prince. 1993. Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes*, 8:1–56.

Roelofs, A. 1992. A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42:107–142.

Rumelhart, D. and J.L. McClelland. 1986. On learning the past tenses of english verbs. In J. McClelland and D. Rumelhart, editors, *Parallel distributed processing*, volume 2. MIT Press, Cambridge, MA.

Seidenberg, M. and M. Bruck. 1990. Consistency effects in the generation of past tense morphology. *Paper presented at the 31st meeting of the Psychonomic Society (November: New Orleans).*

Seidenberg, M.S. 1992. Connectionism without tears. In S. Davis, editor, *Connectionism: Theory and practice*. Oxford University Press, pages 84–122.

Seidenberg, M.S. and J.H. Hoeffner. 1998. Evaluating behavioral and neuroimaging data on past tense processing. *Language*, 74:104–122.

Seidenberg, M.S. and J.L. McClelland. 1989. A distributed, developmental model of word recognition and naming. *Psychological Review*, 96:523–568.

Shen, D. and K. Forster. 1999. Masked phonological priming in reading chinese words depends on the task. *Language and Cognitive Processes*, 14:429–459.

Stanners, R.F., J.J. Neiser, W.P. Hernon, and R. Hall. 1979. Memory representation form morphologically related words. *Journal of Verbal Learning and Verbal Behavior*, 18:399–412.

Stemberger, J.P. 1993. Vowel dominance in overregularization. *Journal of Child Language*, 20:503–521.

Stemberger, J.P. 2004. Phonological priming and irregular past. *Journal of Memory and Language*, 50:82–95.

Stemberger, J.P. and B. MacWhinney. 1988. Are inflected forms stored in the lexicon? In Michael Hammond and Michael Noonan, editors, *Theoretical morphology: Approaches in modern linguistics*. Academic Press, San Diego, pages 101–116.

Stemberger, J.P. and C.M. Middleton. 2003. Vowel dominance and morphological processing. *Language and Cognitive Processes*, 18:369–404.

Talairach, J. and P. Tournoux. 1993. *Referentially oriented cerebral MRI anatomy: Atlas of stereotaxic anaotmical correlations for gray and white matter*. Thieme Medial Publishers, New York.

Ullman, M. 1999. Acceptability rations of regular and irregular past-tense forms: Evidence for a dual-system model of language from word frequency and phonological neighborhood effects. *Language and Cognitive Processes*, 14:47–67.

Ullman, M. 2001. The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, 30:37–69.

Ullman, M., R. Bergida, and K. O'Craven. 1997. Distinct fmri activation patterns for regular and irregular past tense. *NeuroImage*, 5:S549.

Ullman, M., S. Corkin, M. Coppola, G. Hickok, J.H. Growdon, W.J. Koroshetz, and S. Pinker. 1997. A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are preocessed by the procedural system. *Journal of Cognitive Neuroscience*, 9:266–276.

Wagner, A.D., R.L. Buckner, W. Koustaal, D.L. Schacter, J.D. Gabrieli, and B.R. Rosen. 1997. An fmri study of within- and across- task item repetition during semantic classification. *Cognitive Neuroscience Society Annual Meeting*, 4:68.

Wagner, A.D., W. Koustaal, A. Maril, and D.L. Schacter. 2000. Task-specific repetition priming in left inferior prefrontal cortex. *Cerebral Cortex*, 10:1176–1184.

Ward, B.D. 2001. Deconvolution analysis of fmri time series data. *Medical College of Wisconsin*.

Weyerts, H., M. Penke, U. Dohrn, H. Clahsen, and T. Münte. 1997. Brain potentials indicate differences between regular and irregular german plurals. *Neuroreport*, 8:957–962.

Zarahn, E., G. Aguirre, and M. D'Esposito. 1997. A trial-based experimental design for fmri. *NeuroImage*, 6:122–138.