This is to certify that the

thesis entitled

THE EFFECT OF TASK DEMANDS AND STIMULUS VARIATION

ON THE MAGNITUDE OF REPETITION EFFECTS

presented by

Joseph Spencer Brown

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Psychology

TH Carr

Major professor

Date February 18, 1991

THE EFFECT OF TASK DEMANDS AND STIMULUS VARIATION ON THE MAGNITUDE OF REPETITION EFFECTS

By

Joseph Spencer Brown

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Depatment of Psychology

1991

ABSTRACT

THE EFFECT OF TASK DEMANDS AND STIMULUS VARIATION ON THE MAGNITUDE OF REPETITION EFFECTS

By

Joseph Spencer Brown

A central issue in the study of memory is the mechanism by which information is stored and how practice effects occur. A current debate centers around the question of whether information is stored as specific memories of previous encounters (instance theories) or as abstract or general traces (abstractionist theories). Proponents of each view look to the effects of the variation of surface form on repetition effects in letter and word processing. In support of instance theories, an effect of surface form match on repetition often occurs when tasks remain the same from one encounter to the next, while these effects are not found when tasks do not change. In the first of two experiments, the effect of task match, or mismatch, on the magnitude of surface form effects was examined. Sixty-four subjects viewed words and non-words and were instructed either to pronounce the letter string or indicate whether it formed a word. Half of the subjects performed the same task in a second block of trials, while half performed another task. During the second block, subjects saw half the words in the same script (handwritten or typed) as the first block and half in the other script. No effect of task change was found on the magnitude of repetition effects, nor did task change affect the effect of surface form match or

mismatch. Handwritten stimuli were responded to more quickly in both tasks if they had been viewed in the same form on the previous block. Typewritten stimuli were unaffected by the match or mismatch of surface form. In the second experiment, the correlation of subjects memory for surface form and the magnitude of surface form effects was examined. There was no relation between subjects memory for surface form and the magnitude of surface form effects in experiment one.

To Cindy

and

Cliff

ACKNOWLEDGEMENTS

I would like to gratefully acknowledge the members of my committee: J. Katherine Bock, David E. Irwin and Rose T. Zacks. Your patience was a godsend. Special thanks to the chairman of the committee, Thomas H. Carr, whose attention to my education often exceeded my own. Finally, thanks to the faculty of the psychology department of the University of Nebraska-Omaha for their support for the completion of this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

If, as psychologists propose, our behavior is the result of the interaction of our genetic background with our experience, then one of the most important aspects of understanding human behavior is the understanding of how experiences are stored so that they can create changes in behavior. This accounts for the plethora of theories which attempt to explain changes in performance as a result of practice.

Though theories of skill acquisition and improvement can be categorized according to a variety of taxonomies, one of the most important current distinctions is between episodic or instance theories (e.g. Jacoby & Brooks, 1984; Jacoby & Hayman, 1987; Kirsner, Dunn, & Standen , 1987; Kolers, 1976; Roediger & Blaxton, 1987; Logan, 1988), and abstractionist theories (Morton, 1969; Carr & Pollatsek, 1985; Carr, Pollatsek & Posner, 1981). I will use the term 'instance theory' to describe the theories that argue that experience improves performance by allowing us to retrieve specific memories of past performances. These specific memories are used to help guide future performance. The more similar the first experience and the second, the more the first experience improves performance on the second encounter. The term 'episodic theory', though often used to describe these theories, can easily lead to confusion with theories of episodic memory proposed by Tulving (1972, 1983). As will become clear when I describe the evidence that supports instance theories, these are not theories about the autobiographical memory described by Tulving, but rather, they are attempts to describe performance in

tasks that are more commonly classified as the domain of semantic memory: priming of lexical decision, word identification, or word stem completion for example[1]. Though it is possible that the patterns of results in semantic memory tasks will mirror those found in autobiographical memory (perhaps because they rely on the same underlying representation), the question of what variables affect conscious recall is independent of questions about repetition priming. That is, to find that context effects, for example, have an impact on conscious recall says nothing about whether they will have an effect on a more semantic task, for instance repetition priming.

In contrast to instance theories stand abstractionist theories which argue that experience improves performance because we abstract from it general principles or abstract pieces of information that are not tied to any specific instance, but can nevertheless be applied as a guide to future performance.

The goals of this dissertation are threefold. First, it will give an overview of some of the evidence supporting both instance and abstractionist theories, and, by way of clarifying the nature of these theoretical positions, describe some examples of particular theories based on these data. Second, it will apply an organizational framework that makes predictions about the circumstances under which abstractionist results will be found and the circumstances under which data favoring instance theories will emerge. The description of this framework and the proposal of tests of it will be the primary objective of the introductory section. The body of the paper then presents the experiments that constitute the empirical investigation of the organizing framework. Finally, it will attempt to interpret the data obtained in the experiments with respect to the framework, and suggest the direction of future research.

To begin, let us examine some of the best evidence for instance theories of the effects of experience within the domain of the classic repetition effect

paradigm using words or text as stimuli. While the abstractionist models predate instance theories (at least in their current manifestation) I would prefer to ignore the chronology and begin with an outline of instance theories and the data that support them.

Instance accounts of repetition effects in reading are based on evidence from three types of experiments. All of these experiments rely on a change in presentation conditions between the first and second presentation of a word. According to instance theorists, changes in the magnitude of repetition effects as a result of changes in stimulus properties other than word identity between presentations is evidence that recognition is not the result of abstract representations which capture only word identity, but rather, the result of information that represents the specific, concrete details of a particular encounter with a word. Three types of variation that have been particularly fruitful for instance theorists are variation in modality (e.g. Jacoby & Dallas, 1981; Kirsner & Smith, 1974), context (e.g. Jacoby, 1983a, 1983b), and physical features of the letters in which the stimuli are printed such as orientation, type font or case (e.g. Masson, 1986; Kolers, 1973, 1975, 1976; Kirsner, 1973, Jacoby & Hayman,1987).

The effects of changes in modality between exposures to a stimulus are considered by some to be an excellent example of evidence consistent with instance theories (Levy, personal communication). Proponents of this view find the experiments like those of Kirsner and Smith (1974; Kirsner, Milech & Standen, 1983) particularly convincing. In this experiment, subjects were presented with words and non-words in a lexical decision task. Stimuli were presented either auditorily or visually. Of interest was the magnitude of the repetition effect both within and across modalities. Repetition facilitation was

greater for matched modalities than for mismatched modalities regardless of the target modality.

Clearly, this evidence contradicts the strongest possible form of the abstractionist account, which would argue that regardless of stimulus characteristics, the magnitude of repetition facilitation should be equal, so long as the items possess the same identity and hence activate the same abstract representation. However, though this evidence is extremely damaging to the strongest form of the abstractionist position, there exist at least two quite tenable restatements of the account. The first, proposed first by Morton (1979; Jacoby, 1983a), simply argues that there exist different abstract representations for words in different modalities. This stance loses some of the elegance that a single representation has, but it does handle the data without abandoning a substantially abstractionist position. Alternatively, given the considerable separation of the neural structures that subserve audition and vision, it is not unreasonable to claim that the difference in facilitation for matched and mismatched modalities can be found in the more peripheral structures. That is, there may exist a modality independent abstract representation which is activated by the two perceptual systems. This representation may be activated equally whether the stimulus is auditory or visual. However, the perceptual systems responsible for the translation from the proximal stimulus to the abstract representations may demonstrate a sense specific priming effect over and above the repetition benefits associated with the abstract representation.

Though effects of modality on repetition benefit are often found, it is important to note that this is not always the case. For instance, Srinivas and Roediger (1990) fail to find modality effects in primed category associations. Though one might wish to use such failures to find surface form effects as evidence for the abstractionist position, it seems more reasonable to consider a

more moderate view taken by Roediger and his colleagues (Srinivas & Roediger, 1990; Roediger & Blaxton, 1987) which argues that modality and surface form effects will be found only in tasks that rely heavily on bottom up perceptual processes (i.e. data driven tasks). Such a theory nicely describes failures to find modality or effects in tasks such as category association, but, as you will see, seem to have a great deal of problem dealing with failures to find surface form effects in primarily data driven tasks.

The sort of arguments that have been made with respect to modality variation lead us to a larger argument that will run throughout the rest of this paper. Given the sorts of evidence presented in cross modality priming studies, the correct way to consider the abstractionist/instance theory debate is not in terms of whether the strongest form of the abstractionist account can or cannot be sustained. Clearly, the strong form of the abstractionist account falls in the face of the evidence. However, this does not mean that the strong form of the instance theories is necessarily correct. The question 'is there a totally abstract representation?' is replaced by 'do all variations in stimulus characteristics that have an effect on processing also reduce repetition effects?'. In some sense, the abstractionist position is that there exists a representation that is insensitive to some subset of perceptually relevant variation. The term perceptually relevant is used here to denote variations in the stimulus that have a marked effect on the initial processing of the stimulus. It would be unwise to require instance accounts to predict variations in repetition effects on the basis of differences such as small variations in ambient lighting, contrast of the stimulus and its background, etc. Though evidence showing that these trivial variations reduce repetition effects would indeed be impressive, to require such evidence in order to refute the abstractionist account lays too heavy a burden on instance theorists. To see this, recall that it is the similarity of two memory traces that

facilitates practice effects in the instance accounts. Thus, if two stimuli are sufficiently similar, the transfer from one exposure to the second may be extremely effective, producing results that are indistinguishable from complete transfer. Therefore, though the abstractionist position does not, in the most general sense, require that the representations be insensitive to all variation, it does require that the representations be insensitive to a large subset of variations that affect processing. The instance account, on the other hand, must predict that practice effects be sensitive to all variations that affect processing. Obviously, the more types of perceptually relevant variations that are found to be irrelevant to practice effects, the more globally satisfying the abstractionist account is, conversely, the more types of variations that are found to affect practice, the more satisfying the instance account. As I hope to demonstrate, neither theory can, in its most extreme form, account for all the data. The best we can manage is an organizing structure that suggests under what conditions we should expect to find sensitivity or insensitivity to variation.

One sort of stimulus variation that might be ignored in a quite abstract word recognition system is context of occurrence. Context can be taken to be any set of percepts that is processed in close temporal proximity to the stimulus. Thus, changes in the words making up a sentence in which a target word occurs between a first and second exposure of that word constitutes a context change. Likewise, a change in the environment between first and second exposure to a word (for instance changing the room in which first and second exposure occur) constitutes a change in context. Finally, internal percepts such as the mental operation that generates a word or must be performed on a word can provide a context. Context change nicely fits our criterion for a fair test of instance versus abstract theories, since it can be easily shown to affect initial processing. For instance, words in a normal sentence are read more quickly

than a random list of those same words (e.g. Carr , Brown, & Charalambous , 1989). At the same time, the word itself has not been changed by changes in context. As Jacoby (1983a) argues, the abstractionist account of reading predicts that the repetition of a word should result in the same facilitation regardless of the context at either exposure, assuming the word has only one meaning. With this prediction in mind, Jacoby conducted a series of experiments all within the same general framework. In these experiments, subjects were shown three types of practice trials. In the first type of trials, a series of X's was presented, followed by a presentation of the target word. Subjects were told to name the target word as quickly as possible. In the second type of trial, a word was presented, followed by its antonym. The subjects' task was to name the antonym. In the final type of trial, subjects were presented with a word, followed, at an interval, by a series of question marks. The subjects task was to generate the antonym, which becomes the target word. In all three conditions, the target word was the same. In the test condition that is of most interest, subjects were presented with the target word masked by a series of ampersands (&&&&), and were asked to name the target. The dependent measure was the percentage of words correctly identified. Jacoby argues that the abstractionist account should predict that the ability of subjects to name the target in the perceptual identification task will be independent of the practice condition, because the same abstract representation should be activated regardless of the context or the mental operations performed on the stimulus. This was not the case; subjects were most able to name targets that had been previously presented without context, while the naming of targets that had been generated was considerably less accurate, as was the naming of those targets presented within the context of their antonym. These results are supported by a number of studies that produce similar patterns of results, with

naming latency replacing accuracy as the dependent variable (Levy & Begin, 1984; Mason & Freedman, 1985; Carroll & Kirsner, 1982; Monsell & Banich (described in Monsell, 1985)). Across considerable variation in the experimental detail, change of context reduced repetition benefits in each of these studies.

Experiments that find differences in repetition effects as a result of case, type font or orientation changes have also been presented as very strong pieces of evidence for the instance theory account of perceptual recognition. One example of this type is found in a paper by Masson (1986). In this experiment, stimulus words were made up of letters that were written backwards. In addition, the words were presented in mixed case (MiXeD CaSe). In the experiment most damaging to the the abstractionist position, consisted of a practice phase and a test phase. In the practice phase, a groups of three words were presented and subjects were told to pronounce the three words as rapidly as possible. After completing the practice lists, the test phase began. In the test phase, subjects were presented with groups of three words. Each group of three words was either new or it had been seen during the practice phase. Groups that had been seen before were of two types, either matching or not matching the case from the test phase, so that MiXeD- cAsE- wOrDs would be tested with either MiXeD- cAsE- wOrDs, or mIxEd- CaSe- WoRdS. Assuming that the abstractionist account is true, you would expect no difference between groups presented in the same case in test and practice phases and groups presented in different cases. Unfortunately for abstractionist theories, subjects performed no better on old groups presented in a new case than they did on totally new groups.

Additional support for the role of surface variations in the processing of words comes from Jacoby and Hayman (1987). In this set of experiments,

subjects were shown a list of words in one of a variety of surface forms (upper case vs lower case, distorted fonts etc). Each of these words was read aloud as quickly as possible. Then subjects were introduced to a new task in which they were asked to identify a series of words in a masked perceptual recognition procedure. The words in the second task were of three types; new words, old words of the same surface form as the first exposure or old words in a new form. Again, the instance theorists predict that match between surface form of the first exposure and the second exposure should increase repetition benefit, while the abstractionists would predict it should not. While surface form match effects were not found with all stimuli, surface form match made a difference in the magnitude of the repetition effect for targets written in small case. [2]

With evidence of this type supporting the instance account of the benefits of repetition, it is not surprising that this class of theories is enjoying a great deal of popularity. Theorists who rely on this type of data can be divided into two broad categories. One group, best exemplified by theorists such as Kolers (1976; Kolers & Roediger, 1984) or Jacoby and Brooks (1984), argue that the information stored in the performance of all tasks is a concrete instance memory of the task, its procedures and the characteristics of the stimuli upon which the task was carried out. These theorists would tend to take a strong reading of the data for instance representations, arguing that, for **all** significant variations, results that favor an instance account should occur. Kolers, for instance, argues that subjects store the procedures that are used in decoding inverted text as part of the memory of that text. Thus, when faced with reading an inverted text, subjects will read the text more quickly if they had previously read the text in an inverted form than if they had read the same text in normal orientation.

A quite similar view is proposed by Logan (1988, 1990) to account for the acquisition of automaticity in all tasks. In short, Logan argues that the data

associated with the acquisition of automaticity can be accounted for by assuming that the performance of a task can either rely on an algorithm or on recollection of instances from memory. He argues that the time it takes to recall a pertinent memory is a monotonically decreasing function of the number of such memories. Thus, as more memories are accumulated, reaction time improves, and it becomes unnecessary to use capacity demanding algorithms. This sort of theory of automatization makes the same predictions about skill acquisition as instance theories of repetition effects make about the word recognition task, so long as the instances that are encoded consist of the physical features of the event rather than some abstract trace. That is, practice acquired with stimuli of certain physical characteristics will transfer poorly or not at all to stimuli with different characteristics. While Logan does not explicitly maintain that the trace must include surface form information, he does imply this link when he sets his theory up as different from theories using abstraction as an explanation for performance (p.494), and he agrees with this interpretation of his position (Logan, personal communication).

In an attempt to explain similar data regarding context effects in explicit recognition and recall, Tulving and Thompson (1973; Thompson & Tulving,1970) invoke the notion of encoding specificity; that is, the contention that the conditions under which a stimulus is encoded affect the properties of the trace. Thus, a word initially encoded with a certain associate is remembered better if it is recalled in the presence of that associate, rather than in the presence of a normally more effective cue. In addressing subjects' ability to utilize specific autobiographical memories in explicit memory tasks, Tulving and Thompson are clearly addressing a different set of tasks than are theorists attempting to explain context effects in perceptual word recognition. In fact, Tulving (1983) argues that the operating principles found in explicit episodic

memory do not apply to tasks of the later type. However, it is equally clear that instance theorists like Kolers and Jacoby and Brooks would like to extend the logic of Tulving into this new task domain, arguing that all performance rests on the use of the same sorts of specific episodic memories.

Other instance theorists take a more cautious stand on the issue of abstract versus concrete memory representations that support practice. These theorists (Roediger & Blaxton, 1987; Morris, Bransford & Franks,1977) argue that under some task conditions concrete memories are stored and used in later processing, while under other conditions more abstract representations are encoded. For instance, Roediger and Blaxton (1987) argue that most implicit memory tasks, such as perceptual identification, lexical decision or, in their view, fragment completion, rely heavily on perceptually driven processes and are therefore likely to provide evidence consistent with instance theories of performance. Explicit memory tasks, such as free recall or imagery rely more heavily on conceptually driven processes and are therefore more likely to provide evidence for abstract storage in memory. Thus, they argue:

" ... most implicit memory tasks in their standard form depend on data driven processing; they should be sensitive to the match of surface characteristics between study and test....On the other hand, most explicit tests benefit from conceptual processes (elaboration, generation, imagery, etc.) and are little affected by altering surface characteristics." (p. 386)

In spite of the current popularity of instance theories, there exists evidence that contradicts their predictions. Remember that instance theories predict that differences in surface form and context between first and second exposures of a target should affect the amount of facilitation due to repetition, at least within perceptually driven tasks such as word recognition. Failure to find such effects is good evidence for abstractionist theories of reading. These

failures are fairly common, especially with respect to changes in surface form (Levy, Newell, Snyder & Timmins, 1986; Brown, Sharma, & Kirsner, 1984; Rayner, McConkie & Zola, 1980; Scarborough, Cortese & Scarborough 1977; Carr, Brown, & Charalambous, 1989; and (with some targets) Jacoby & Haymen, 1987), but also exist with changes in context (Dixon and Rothkoff,1979; Jacoby 1983b (with respect to environmental context); Carr, Brown, & Charalambous, 1989). In order to give a flavor of these findings, I will give you an example of each type of failure.

An example of failure to find effects of surface form is found in Levy et al's (1986) study of proofreading. In this study, subjects repeatedly proofread passages. These passages were either all in the same type font or in a different font on each repetition. Reading times and error rates were the same regardless of the font match or mismatch. Over the course of four exposures, reading times improved regardless of font match.

An instance of failure to find context effects is found in Levy (1983). Here, subjects were instructed to proofread passages for spelling errors either with or without prior familiarization with the passages. The target passages were either identical to the passages which subjects studied or scrambled versions of these paragraphs; that is, the same words in a new context. Repetition facilitation, as measured be the number of errors found in a limited time, was the same whether the context matched or mismatched.

One might argue that proofreading is a more conceptually driven task, and thus unlikely to produce evidence of concrete perceptual representations. However, Carr et al (1989) also failed to find the effects predicted by instance theorists in a pronunciation task, a task that in the Roediger and Blaxton framework is a largely perceptually driven task. In these experiments, subjects were asked to read prose passages aloud. In a series of four experiments, the

effects of constancy of context and surface form were examined. To examine the effects of context on word recognition, subjects were asked to read passages which were either normal prose passages, or passages containing the same words in the same font but in a new randomly scrambled order, placing each word in a new context. Thus subjects would read two passages of the same words, but those words would either be in the same order on each repetition, or in two different orders. Contrary to instance theories, reading times for the second passage in each pair were independent of the match or mismatch of context between that passage and the version that had been read first.

To test the importance of the surface form of the texts in repetition effects, subjects were presented with pairs of texts consisting of the same words in the same order, but either matching in surface form or mismatching. Each passage could either be typed or handwritten. Thus, subjects read the same words twice, either in the same form or in two different forms. Reading times for the second texts were independent of the form of the first text. Therefore, in contrast to the predictions of the instance theorists, match or mismatch of physical appearance of texts had no effect on the magnitude of the repetition effect. In other experiments, this lack of an effect of surface form was found to exist regardless of subjects expectations about surface form and to hold true with orthographically regular non-words as well as words.

The Carr et al (1989) results are especially problematic to the more radical instantiations of the instance theory position as they present what is in many respects the paradigmatic case where surface form effects should arise. That is, physical differences between the two surface forms were great, and in addition, could be demonstrated to have affected processing. The effects on processing can be seen in slower reading speed for handwritten stimuli relative

to typed stimuli and in greater repetition effects for handwritten texts than typed texts. Further, in one of the experiments Carr et al conducted, the stimuli were random sets of words making the task as close to purely perceptually driven (at least in Roediger and Blaxton's framework) as is possible. These data are especially difficult for Kolers variation of instance theory to deal with, as the two surface forms can be shown to be processed differently, as they respond differentially to practice. However, any version of an instance theory has a very difficult time with failures to find surface form effects. To see this, consider the logic of the instance theory position as eloquently stated by Jacoby and Hayman (1987, p. 456):

> A prior presentation of a word can have a large and long-lasting influence on its later identification (e.g. Jacoby, 1983a; Jacoby & Dallas, 1981). Does this effect of experience depend on preserving the visual details of a word, such as type face, between its prior presentation and later test? By a popular view of perception, variation in visual details should be unimportant. Word identification is treated as depending on the identification of abstract letter units, with information about visual details being lost very early in the reading process . . . In contrast, we (e.g. Jacoby, 1983a, Jacoby and Brooks,1984) have suggested that perception can rely on memory for prior episodes. That is, we have argued that identification of words does not totally rest on the use of some abstract representation. Rather, memory for a prior encounter with a word can be retrieved and can serve to aid its later identification. . . By this view, changing the visual details of a word between its presentation and test can reduce transfer by decreasing the similarity of the operations used to identify the word on the two occasions. In light of these claims, it is important to demonstrate that effects of prior experience in later word identification can reflect memory for supposedly superficial details of that prior experience.

As you can see, instance theorists consider surface form match effects to be of great importance. Although you can consider the instance/abstractionist debate with respect to any perceptually relevant variation across stimulus presentations, surface form is perhaps the most dramatic arena in which to contrast these theories. After all, because a variation in surface form has no effect on the identity or meaning of word, an abstractionist theory must predict an absence of surface form match effects, while an instance theory, faced with the obvious physical differences, must predict these differences. For this reason, the remainder of this discussion will emphasize surface form match effects as a domain in which to examine these two positions.

At this point then, it seems we have reached a theoretical impasse. With two sets of experiments that presumably measure the same thing finding different results, how can we evaluate the theories they are supposed to test? There are several possible explanations for these contradictions, but they all boil down to the fact that the experiments are not, in fact, equivalent.

To see that the experiments that favor instance theories are not equivalent to the experiments that find results consistent with abstractionist theories, let's consider the experiment by Masson (1986) that I described earlier. Note the difference between this experiment and the experiments reported by Levy et al and Carr et al: Masson presented his mixed case words in what amounted to a completely new orthography (mirror image texts) while the others used orthographies with which the subjects were extremely familiar. In the same vein, Kolers (1973,1975,1976) found his surface form dependent effects using inverted scripts, hardly a standard orthography. Similar differences are noted by Jacoby and Hayman (1987) who agree that experiments in which the target texts are in novel orthographies can truly only speak to the contribution of instance memories in the early stages of learning,

conceding that the true test of surface form effects is in variations among well learned scripts.

As a somewhat exaggerated example of why this is the case, consider learning a new orthography for English in which X and Y represent the same phoneme. An extreme view of the abstractionist account might argue that as long as the reader can perform the translation from script to phoneme, there should be no surface form effects, since the same phoneme is activated in both cases. An abstractionist account of this type is easily refuted by the evidence. However, such an account makes the assumption that the only possible locus of practice effects is at the phoneme level or higher, and that lower levels do not require any learning in order to translate both X and Y to the abstract phonological code. Such assumptions are simply too restrictive. Though it is true that at high levels of learning the abstractionist account must predict an absence of surface form effects, it does not predict them with a new script.

Seen in this light, it is hard to imagine a set of experiments that are less likely to find compelling evidence of abstract representations than those that used novel orthographies. In order for the system to utilize abstract representations of a series of words, it must first have some experience with the range of surface variation associated with typographies of these words. For instance, subjects have learned that K and k both are representatives of the category 'cay'. With a new orthography subjects simply don't have the information needed to define the boundaries of the category. The more instances of the category the subject sees and the greater the variability of these instances, the more able the subject is to identify new exemplars of the category (Posner & Keele,1968; Posner, Goldsmith & Welton, 1967). In experiments in which the stimuli consist of words written in a new orthography,

subjects have not had sufficient experience with the stimuli to form and make use of the abstract representations of the category which the stimuli represent.

Though novel orthographies can be called upon to account for many of the effects found in the repetition literature that are consistent with the instance theory account, they cannot deal with all such effects. However, another difference can be found between those tasks that find instance influenced effects and those that do not. The second crucial difference between experiments that find instance effects, for example Jacoby (1983) or Jacoby and Hayman (1987), and the ones such as Levy et al (1986) or Carr et al. (1989) which find abstractionist effects, is that in the abstractionist experiments subjects performed the same task in both the practice and test trials, whereas in the experiments that find instance driven effects subjects performed two different tasks. In both Jacoby and Hayman (1987) and Jacoby (1983b) speeded naming was the task on the first exposure and tachistoscopic full report was the task on the second. The difference in these tasks, while possibly subtle from some perspectives, is more evident after further consideration; there is a difference, at least at a surface level, between a speeded naming task and a perceptual identification task in which speed is not as highly stressed, but in which the target is degraded by a mask. It is not unreasonable to suppose that the mental operations involved in these two tasks are quite different, for instance, tachistoscopic identification tasks of the type described are more prone to guessing strategies than a naming task in which adequate viewing time is allowed (Smith & Spoehr, 1974; Schvaneveldt & McDonald, 1981; Ratcliff, McKoon, & Verwood, 1989;).

The difference between the experiments that find surface form affects and those that do not is instructive; when the task remains constant between first and second repetition, the context becomes irrelevant (Carr et al), but when the

task changes between repetitions, the context in which the target was originally presented becomes important (Jacoby). This pattern of results illuminates similar findings about surface form effects in repetition benefits reported by Kirsner, Dunn, and Standen (1987) in which a lexical decision task was used in the practice phase and a perceptual identification task was used in the test phase. Other studies which find effects of surface form and context changes also vary task demands between first and second exposure (Roediger & Blaxton,1987; Jacoby & Hayman 1987; Levy & Begin, 1984; Levy, 1983; Jacoby & Witherspoon, 1982). In a recent review, Carr et al (1989) argue that this pattern of results is true throughout the literature; those studies that find surface form effects with normal scripts also change task between exposures. Though they ultimately concentrate on surface effects alone, independent of task constancy or change, Jacoby and Brooks (1984) suggested a similar possibility stating that the role of surface characteristics in memory will depend on the prior processing conditions and their match with processing conditions at retrieval.

At this point it is important to note again that the theoretical position that is suggested by the emerging pattern of results is neither entirely consistent with abstractionists, who can't explain why surface form match and context sometimes effect performance, nor is it entirely consistent with instance theorists, who can't explain failures to find effects of surface form or context; neither theoretical position can stand up to the data at hand. The position that I believe emerges is one that argues for a system that produces data consistent with abstractionist theories so long as task remains consistent across stimulus exposures, but produces data consistent with instance theories when task changes.

This mixing of the abstract and instance based representation is not all that radical a departure from the views of abstract theorists such as Morton.

Even Morton, that most abstract of abstractionists, concedes that there must be separate representations for spoken language and visual language (Morton, 1979) in order to account for the reduction of priming effects across modalities. A similar sort of argument could account for the apparent interaction of task change and surface form effects found in the literature. All you need to do is hypothesize a separate representation for each task that is abstract with respect to surface form and context, and which is accessible only to the task in which it is originally encoded. The problem with this sort of explanation is that it requires a multitude of representations thus losing much of the appeal of abstractionist theories. In addition, this sort of explanation has a very hard time with new tasks such as lexical decision for which the subjects could not have an abstract task representation prior to the experiment. Fortunately, I do not believe it is necessary to propose a totally independent representation for each task. Instead, I would propose that the same set of representations are available in each task situation, but these representations are retrieved differently, depending on the task conditions across exposures.

In order to see how a single type of representation might show both instance based or abstractionist data, consider the types of retrieval cues that are available to retrieve past experience when confronted with a new stimulus. In addition to the context of occurrence and the surface form of the stimulus, there are also the mental operations involved in the original storage of the stimulus. If the surface form of the stimulus is used as a retrieval cue, then there should be greater ease of retrieval if surface form remains constant across exposures and thus, surface form match would effect repetition benefit. If, on the other hand, the mental operations are used as a retrieval cue, then so long as task remains constant, there will be no effect of changes in surface form across repetitions. One coherent way to explain the pattern of results relating surface

form match effects to task constancy is to argue that the mental operations that occur when the first trace is stored are the primary retrieval cues used on the second exposure so long as those cues are available. Thus, if these operations remain the same, no effects of surface form match are found. When these mental operations are not available as cues because of changes in the task, the secondary cue, surface form is used, resulting in the surface form match effects.

Why might one find that task is a more important retrieval cue than surface form? Two explanations are possible. First, consider the number of cues associated with the performance of a task. While surface form might be considered as one cue, the number of mental operations that are involved in completing a task such as lexical decision or masked perceptual recognition is certainly higher than one. Thus, if each operation can serve as a cue, then there are relatively more cues associated with task than with surface form, thus they might prove to be more important by sheer weight of numbers. Alternatively, it is possible that the cues that are used for retrieval are hierarchically organized with the mental operations simply taking a higher priority than surface form cues, when both are available.

The two differences between experiments that find surface form and context effects on repetition effects and those that do not, suggest an organizational framework which can be used to predict the differences in their results. First, when different mental operations are performed on a word or text in both practice and test phases, the surface form and context of the text become relevant, even when the word or text is in a familiar orthography. I will refer to this as the 'task change hypothesis'. Second, when subjects are faced with a radically new orthography, the match or mismatch of surface form has an impact on repetition effects. I will call this the 'new orthography hypothesis'. In

the rest of the paper, I will be concentrating primarily on the task change hypothesis.

Though the arguments surrounding the task change hypothesis sound quite similar to those of the more moderate instance theorists, such as Roediger and Blaxton or Kolers, discussed earlier, there are important differences to be kept in mind. Note, for instance, that Kolers predicts that any difference in the operations performed between first and second repetition should result in a lessening of the repetition effect. Thus he would predict that changes in surface form without changes in task should still result in a reduced practice effect. In fact he makes just such predictions when he works with inverted texts.

Although there is some evidence to support the claims of Kolers, this evidence not as convincing as one might hope. For instance, the best known evidence for this position comes from transfer studies using a variety of transformed texts (Kolers & Ostry, 1974; Kolers 1973,1975). These studies suffer from the problems with using unfamiliar scripts that were discussed earlier. That is, while they tell us a great deal about the properties of the system during early learning of an orthography, they can tell us little about normal operations.

Other work used to buttress the strong form of Koler's procedural account consists of data relating to subjects' ability to explicitly recognize sentences as having been seen before (e.g. Kirsner 1973; Craik and Kirsner, 1974). As discussed before, such work is potentially relevant, but does not directly bear on the issue of the system's ability to obtain repetition benefits in a performance paradigm. This fact is brought home quite nicely in Kolers (1975,1976) when dissociations are demonstrated between subjects' ability to recognize a stimulus as a previously viewed item, and their ability to benefit from the previous exposure. Such dissociations between explicit memory and

performance are quite common (Jacoby & Witherspoon, 1982; Jacoby & Dallas, 1981; Moscovitch, 1981; Moscovitch, Winocur & McLachlan 1986).

Finally, evidence supporting Kolers position comes from subjects' performance when stimuli are presented in different modalities. As I discussed earlier, such data, while relevant for dismissing one of the more extreme versions of the abstractionist account, are not evidence for instance accounts (and therefore Koler's type of procedural accounts) in general.

The best evidence for the the type of effects predicted by Kolers is found in Monsell (1985) where subjects are tested on their ability to classify words along different sets of dimensions. Subject were asked to classify words based on a syntactic category (nouns versus adjectives) or semantic category (pleasant vs unpleasant). Repetition effects are largest in conditions under which subjects classify words along the same dimension at each exposure rather than along different dimensions. In other work, Ratcliff, Hockley and McKoon (1985) demonstrate a similar pattern of results when subjects first encounter stimuli in an explicit memory task and later are measured in a lexical decision task. While this sort of evidence is a convincing existence proof of the notion that **some** task variations effect the extent of repetition priming, it is a fairly small rock upon which to build the notion that **any** changes in procedure should mitigate repetition effects.

The the task change hypothesis also bears some resemblance to models proposed by Roediger and Blaxton when they argue for the representation left in memory by task performance being a function of task demands. The difference lies in the locus of the surface form effects. Roediger and Blaxton argue that the effect is found in relationship of the original encoding strategies and the current task demands. That is, if on the first encounter with a stimulus, surface form information is stored because of the demands of the task (that is, if

it is data driven), then surface form effects will be found. Thus, the original task demands influence whether or not information about surface form is stored, and should subsequent tasks require surface form information, then surface form effects will be found. However, if the first exposure to stimulus is in the context of a task that does not preserve surface form (i.e. a conceptually driven task), then surface form effects will not be found, regardless of the second task. This sort of conceptual framework cannot account for Carr et al's evidence for abstractionist data found in the context of a perceptual task. Though the Roediger framework does address the issue of the task match/mismatch, at least in predicting surface form effects only when both tasks are data driven, it is quite different than the task change hypothesis that I am proposing. To see this, note that the Roediger framework predicts surface form effects when a data driven task is used at both and second exposure. The task change hypothesis predicts no such effects. The task change hypothesis, on the other hand, predicts surface form effects when two different tasks are used. Though the task change hypothesis does not specifically address that data/conceptually driven dichotomy of Roediger, it is interesting to note that the data on which it is based on tasks that are primarily data driven. Therefore, it may be that there are no surface form effects on conceptually driven tasks as Roediger suggests, but that in order to predict surface form effects in data driven tasks, a different treatment is needed.

There are two caveats that need to be considered before accepting the task change hypothesis whole heartedly. First, though it is true that all of the experiments that find surface form effects also change tasks between presentations, the converse is not true. That is, there are are several studies that find no surface form effects while changing tasks (Levy et al, 1986; Morton, 1979). Secondly, while the pattern holds perfectly for surface form, there are

two exceptions with regard to context effects (Masson and Freedman, 1985; Carrol and Kirsner, 1982). In spite of these concerns, it seems reasonable to test this pattern, at least with respect to surface form. That is, the occurrence of perceptual abstraction, as defined by transfer of repetition benefits across changes in surface form will be controlled by task change. When the task (and hence the mental operations required) stay the same, perceptual abstraction will occur. When the mental operations required for performance are changed, perceptual abstraction will diminish.

An alternative to the task change hypothesis is suggested by examining the types of task that yield surface form effects. After eliminating those which use a distorted or novel script, there remain five clear cases of surface form change effects: Jacoby and Hayman (Experiment 1), Kirsner, Dunn and Standen (1987), Roediger and Blaxton (1987), Jacoby and Witherspoon (1982); and Levy (1983)[3]. The surface form effects found by Levy (1983) do not occur in subsequent work using a similar task (Levy et al., 1986). Of these remaining four cases, none used the pronunciation task or the lexical decision task as a test of repetition advantage for word recognition. In the case of Jacoby and Hayman, Kirsner et al and Jacoby and Witherspoon, the test task was tachistoscopic report. That is, subjects were briefly presented with degraded versions of the target words and asked to identify them. Roediger and Blaxton used a fragment completion task to measure performance. It is quite possible that the there is something special about the tasks that these researchers have chosen that tends to encourage the emergence of surface form effects.

At first reading, this alternative hypothesis might also sound a great deal like the arguments of Roediger and Blaxton, and in a sense it is. However, a fundamental difference lies in the Roediger and Blaxton's attempts to predict which tasks will give evidence of surface form effects. Roediger and Blaxton

make the prediction that so called 'data driven tasks', which include word fragment completion as well as word recognition, should find surface form effects, and more 'conceptually driven' tasks such as free recall, elaboration, etc, will not. Given the evidence collected by Carr et al, it seems clear that not all 'data driven' tasks generate surface form effects.

While the Roediger and Blaxton prediction about the effect of task on surface form effects seems to be incorrect, another possible way of classifying tasks that produce surface form effects is still available. One possible difference between the types of tasks that find surface form effects and those that do not, relates back to the tasks used by Jacoby (1983a). Those experiments that find surface form effects find them with tasks that may be susceptible to guessing strategies, such as masked recognition or fragment completion. Therefore, it is seems possible that the surface form effects arise from subjects' guessing strategies given incomplete information. That is, subjects are more likely to guess within surface forms than across them. In one study of semantic context, Schvaneveldt and MacDonald (1981) found that, in a lexical decision task, context affected subjects' criterion bias when the targets were masked, but appeared only to affect subjects' sensitivity when the target was continuously available. Thus, at least in lexical decision, the presence of a mask leads to effects related to subjects' strategies in performing the task.

This line of thought must remain speculative because the crucial information to support it would be found in an analysis of subjects incorrect responses. If examination of subjects incorrect responses showed that a higher proportion of these incorrect guesses consisted of items that were previously presented in the same surface form as the target, then it would be convincing evidence indeed. Such an analysis is not included in any of these papers, nor is it to be found anywhere. However, the error rates (ranging from 10 to 50

percent) in these studies suggest that subjects are guessing a considerable proportion of the time. Contrast this with typical error rates in lexical decision and pronunciation of high frequency words and you can build a quite good case that if guessing strategies can effect the outcome of experiments such as these, they will do so in the experiments that find surface form effects. Further, it is interesting to note that in Jacoby and Hayman's (1987) experiments, surface form effects appear only when accuracy is low; when accuracy rises to 90% with upper case targets, surface form effects disappear.

To this point, a great deal has been made of failures to find effects of surface form and context in performance data. Of course, the possibility exists that the failures to find these effects are due to lack of experimental power. For instance, Carr et al' s failure to find surface form effects could result from lack of power. Because the subjects were asked to read a series of words and their combined reading times were measured, Carr et al could simply have have had too much error variance to find the effects. The error terms in Carr et al's data indicate that Carr et al had the ability to detect effects in the range of 6 seconds, or about 10% of the total reaction time of 64 secs. Thus, it is possible that there were surface form match effects in Carr et al that were simply not detectable.

Regardless of one's speculations about the underlying cause, the data themselves suggest experiments to test the pattern seen in previous studies. If the task change explanation I have proposed is correct, then we should be able to make surface form variations affect repetition effects when the practice and test tasks are different, and make these effects disappear when the tasks remain the same. With this in mind, I conducted the following experiment.

EXPERIMENT 1

Experiment 1 is a straightforward test of the organizational framework proposed earlier: the task change hypothesis. In this experiment, both task and surface form were manipulated within the repetition effect paradigm. Thus, subjects performed one or both of two tasks: pronunciation and lexical decision. They either performed the same task twice, or performed two different tasks. In addition, stimuli were either shown with the same surface characteristics on the first and second exposure or with different characteristics.

One issue of special importance in evaluating this experiment is the selection of the two tasks in which subjects participate, The tasks must differ from one another in ways which are relevant to testing the hypothesis. What differences are relevant? As the hypothesis is phrased by Carr et. al. (1989), just about any demonstrable difference should be sufficient, so long as both tasks benefit from repetition. This is because the hypothesis is simply a rough summary of the empirical pattern observable in the literature. When the task is identical from one encounter with a stimulus to the next, abstractionist results can be found. When the tasks differ in any way, instance results occur. Pairings of tasks producing instance results include reading and tachistoscopic recognition (Jacoby, 1983a; Jacoby & Haymen, 1987; Kirsner et al. 1987), reading followed by proofreading (Levy et al, 1986), reading followed by fragment completion (Roediger & Blaxton, 1987), and study for a memory test followed by reading (Moscovitch et al., 1986). These task combinations do not lead readily to any particular proposal about a specific process or operation that must be shared between tasks in order to achieve abstraction or must be missing from one of the tasks in order to prevent it. As a result, the task change hypothesis at this point in its development is phrased in very general terms.

For this reason, a straightforward test of the strongest or most general form of the task change hypothesis can be constructed by picking any two tasks that 1) show repetition effects, since insensitivity to practice effects would make them inappropriate on a priori grounds, 2) differ in terms of logical task analysis, such that they appear on logical grounds to require different information for their performance, 3) differ on empirical grounds, in such a way that at least some manipulations produce different effects in the two tasks, 4) can plausibly be defended as data driven, so that even a weak instance theory such as that of Roediger and Blaxton would expect surface form effects. In addition, if tasks meeting these criteria can be found that have already been studied in the literature, the available data will aid in interpreting the results of the new experiment conducted here. Finally, if the tasks have been subjected to formal modeling attempts, then the interpretation of the results in terms of a more specific analysis of shared and unique mental operations might be possible after the fact, allowing the locus of repetition and surface form match effects to be identified, at least tentatively. Two tasks meeting these criteria are naming and lexical decision.

On both logical task analytic grounds and the available attempts to model the tasks, naming ultimately requires phonological and articulatory recoding, whereas lexical decision requires semantic access, a familiarity judgment, or both, at least so long as the nonwords used in the task are orthographically regular and hence pronounceable (Balota & Chumbley, 1984; Carr, Posner, Pollatsek & Snyder, 1979; Coltheart, Besner, Jonasson & Davelar, 1979; Seidenberg & McClelland, 1989). On empirical grounds, there are a number of variables that exert different effects in the two tasks. Lexical decision shows larger word frequency effects and benefits more from repetition than does naming, though both tasks do show effects of both frequency and repetition

(Scarborough, Cortese & Scarborough, 1977). Scarborough, Gerard, and Cortese (1979) report that practice at one of these tasks does not improve performance on the other as much as it benefits performance on the task originally performed, Balota and Lorch (1986) find mediated priming for naming but not lexical decision, and Logan (1988, 1990) reports that practicing lexical decision does not transfer at all to pronounceability judgements. Though judging pronounceability is not the same as actually pronouncing the words, however, the combination of these results and those of Scarborough et al. make one reasonably comfortable with the contention that the two tasks are different from one another. Next, Paterson & Marcel (1977) report that certain brain lesion patients who are unable to pronounce pseudowords perform quite well on lexical decision tasks. This result demonstrates that rule-governed and analogy-based phonological recoding (of the type that would be required for the pronunciation of words) is not necessary for lexical decision, and it is consistent (though it does not demand) theories of lexical decision in which phonological recoding is not involved at all (see Coltheart, et al., 1979; Seidenberg & McClelland, 1989). Finally, both naming and lexical decision have been used many times as tasks in which to observe data-driven perceptual encoding operations related to reading, though it is now generally accepted that lexical decision is more subject to post perceptual decision biases and context effects than is naming (Balota and Chumbley, 1984; Scarbourough, et al. 1977; West & Stanovich, 1982; Seidenburg, Waters, Sanders & Langer, 1984). Putting all this together leads to the conclusion that the naming and lexical decision can serve the purpose of the first test of the task change hypothesis in its strongest form.

One area of concern in evaluating the two tasks that were chosen might be that while they differ in several critical ways, the surface form sensitive

aspects of the tasks may be held in common. Thus, on some theories of the two tasks (see West & Stanovich, 1982; Neely, in press), the pre-access processes that subserve them may be quite similar. Thus, if the surface form match effects are not found, it may because of the similarities rather than the differences between the two tasks. Though such an argument might be used to save a weaker version of the task change hypothesis should the hypothesized effects fail to occur, these tasks still are prime candidates for testing the hypothesis for several reasons. First, the task change hypothesis at least in its current form argues that any task change should encourage the emergence of surface form effects. Second, given our lack of models for such tasks as masked recognition and word completion, it is impossible to argue convincingly for the emergence of surface form effects in any single component of these tasks. The fact that we have models of these two tasks separates them from the tasks upon whose data the theory is built in that our theories of the tasks allows us to point to a locus of the effects should they occur. Finally, because of the amount of study devoted to these two tasks, we at least have considerable empirical data demonstrating that they are different at least on some level, thus we are assured of a test of the task change hypothesis, at least in its current form.

Method

Subjects

Subjects were 64 undergraduates given course credit for their participation.

Stimuli and Procedures

Each subject participated in one session lasting 60 minutes. During a practice period, subjects were shown a block of 10 words and non-words. Depending on what task they were to perform in the first test block, subjects

were instructed to either pronounce each series of letters as quickly as possible while still pronouncing them correctly or to indicate whether each string of letters formed a common English word. It was stressed that this was not a vocabulary test, so there would be no words that were not familiar to them. Another warm up block was run prior to the second test block, with the task performed matching the second task.

In the warm up blocks, and all subsequent blocks, stimuli were presented on a MacIntosh Plus microcomputer. Subjects began each trial by pushing the return key. Subjects were first presented with a warning signal (a '+' symbol) for 1000 msecs. After a 500 msec pause, subjects were shown the stimulus, written black on white, for 500 msec, followed by a blank field. Responses were collected using a Gerbrands G1341 voice key or the MacIntosh keyboard for the naming and lexical decision tasks respectively.

After completing the practice block, the subject completed two experimental blocks. The design for Experiment 1 is found in Figure 1. The experimental blocks represented one cell of the 2 x 2 factorial combination of task (lexical decision or naming) and position (first task or second task). In the first of the two tasks (the practice block), subjects saw 40 words and 40 non-words exposed for 1 second each and were instructed to either pronounce the string or make a lexical decision. Half the strings were handwritten and half typed. Response times were measured in msecs. In the second task (the target block), the subjects were presented with a series of 80 words and 80 non-words and asked to either pronounce the string or make a lexical decision. Half of the strings were new strings and half were seen in the practice block. Surface form (handwritten vs typed) was crossed with previous exposure so that half the old and new words were in each form. Half the old words matched the form of the first exposure and half did not.

First Task	Second Task	Stimuli type 1st Task	2nd Task
Lexical Decision	Lexical Decision		
		Hand	Hand
		Typed	Typed
		Hand	Typed
		Typed	Hand
		*	Hand
		*	Typed
Naming	Naming		
		Hand	Hand
		Typed	Typed
		Hand	Typed
		Typed	Hand
		*	Hand
		*	Typed
Lexical Decision	Naming		
		Hand	Hand
		Typed	Typed
		Hand	Typed
		Typed	Hand
		*	Hand
		*	Typed
Naming	Lexical Decision		
		Hand	Hand
		Typed	Typed
		Hand	Typed
		Typed	Hand
		*	Hand
		*	Typed

(Hand=Handwritten, Typed=Typed, *= Not seen)

Each Subject serves in one of the four cells depicted above.

Figure 1:Experimental Design for Experiment 1.

Stimulus lists were constructed using high frequency words from Kucera and Francis (1967), with word frequencies ranging from 120 to 960 occurrences per million. Across each group of 16 subjects defined by the combination of first and second task, each word was presented equally often in each surface form (handwritten vs typed), in each surface form match condition (same vs different) and as a new or old word. In order to do this, eight stimulus lists of 80 words and 80 non-words were constructed. These eight lists were used in two different orders for each of the four between subject conditions of the experiment. Pronounceable nonwords were constructed by replacing vowels and consonants in 80 words of the same average frequency as the word stimuli.

A list of both words and non-words is included in the appendix. The typed stimuli consisted of New York type font, while the handwritten stimuli were constructed by means of a digitizing tablet using the handwriting of two different writers. These two writers' samples were distributed randomly throughout all conditions.

Analysis.

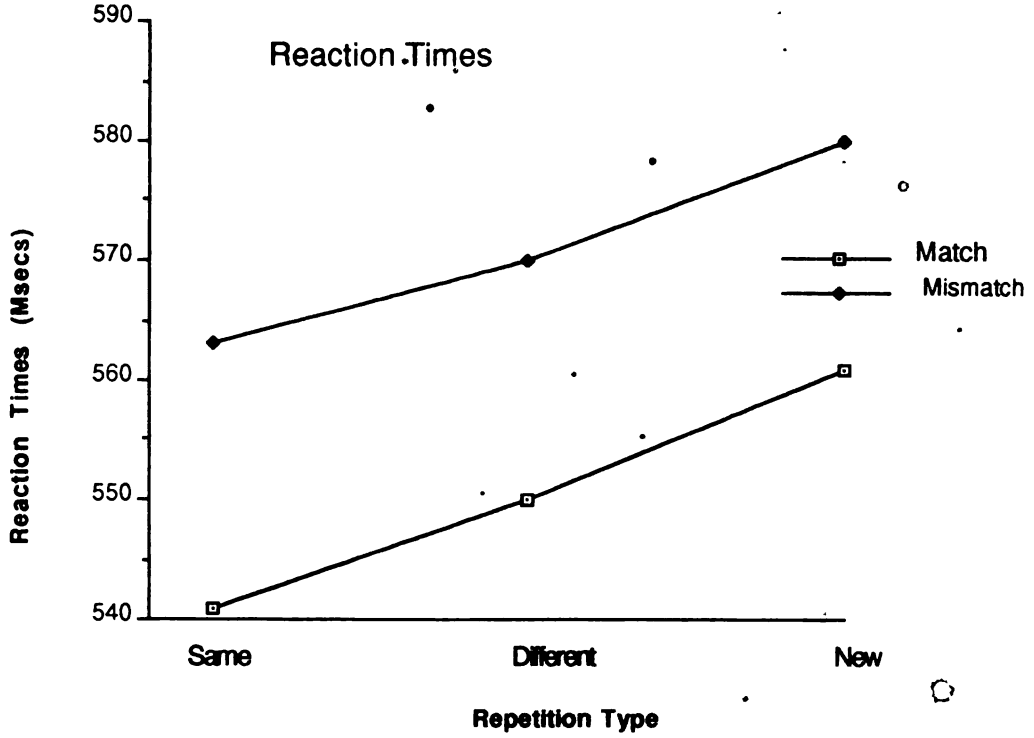
The main analyses consisted of 2 x 2 x 2 x 2 x 3 ANOVA's of the response time for correct responses and error rates for items in the test (or second) block. Errors in the naming task were those trials in which the subject produced a pronunciation that violated the rules of English pronunciation. The between subjects variables were target task (naming vs lexical decision) and task match (first and second tasks match or mismatch), while the within subject variables were surface form (handwritten vs typed), target type (word vs non-words) and repetition type (same surface form, new surface form or new target).

Reaction times of greater than 2000 msecs and less than 200 msecs were discarded.

Results

If the predictions made in the Introduction are true then we would expect an interaction between task match, and repetition type caused by surface form match having an effect on response time and/or error rates when the first and second tasks are different, but no effect when they are the same. Unfortunately for that hypothesis, the task match by repetition interaction is far from significant in both reaction time ($F(2,120)<1$) and error ($F(2,120)<1$) measures. Means for these effects can be found in Figure 2.

Even though the predicted effect was not obtained, the results of the experiment are still quite interesting. For the purposes of the issues addressed in this paper, the effects of interest are those which show the effect of a variable on repetition effects. Before examining these effects, however, I would like to include a brief discussion of several main effects that can be used to validate the method. (a complete list of those interactions that do not involve repetition type are included in Table 1). First, consider the main effect of words vs pseudo words. A main effect was observed such that reaction time to words was faster than non-words in both naming and lexical decision. Next, there was a significant effect of script (handwritten vs typed) such that handwritten stimuli were responded to more slowly than typed stimuli. Finally, repetition benefits were significant in both naming and lexical decsion tasks (though these benefits are not straightforward and are discussed in some detail in the interactions). In addition to a number of main effects and two factor interactions (which are uninterpretable because they are involved in higher order interactions), there



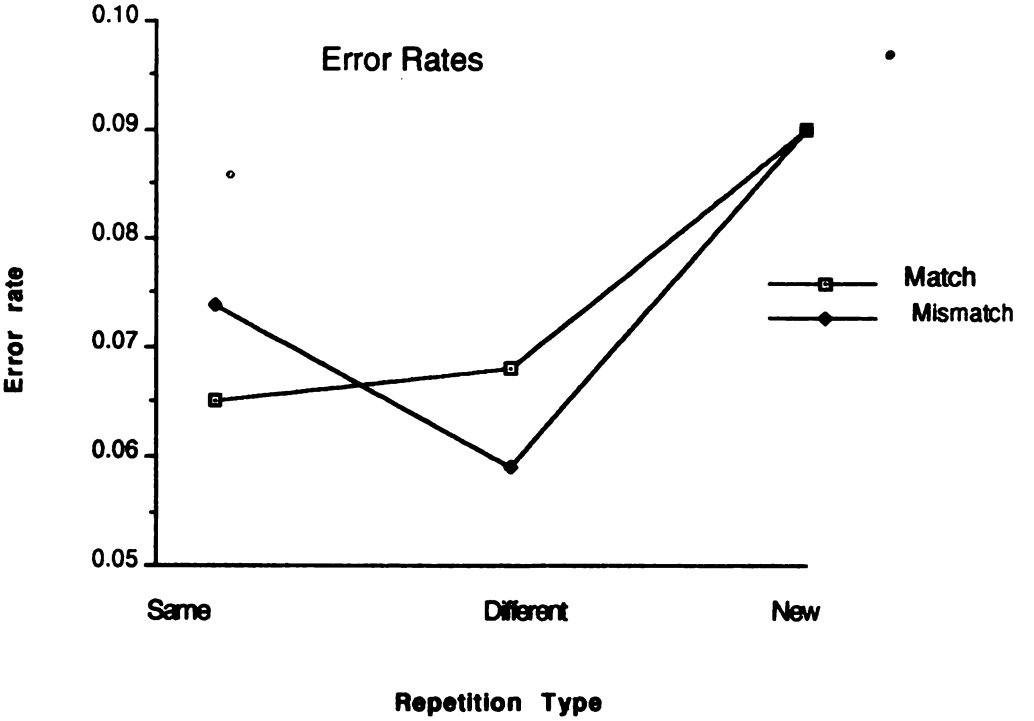


Figure 2: Repetition x Task Match

Table 1: Significant Effects From Experiment 1.

Reaction Times

Source	Df	F	MS_{error}
Task	(1,60)	8.45	5838
Script	(1,60)	42.25	1818
Word	(1,60)	377.39	2732
Repetition	(2,120)	23.11	925
TaskxMatch	(1,60)	4.34	5838
TxW	(1,60)	13.32	2732
TxSxW	(1,60)	16.03	591
TxMxWxS	(1,60)	3.84	591

Error Rates

Source	Df	F	MS_{error}
Script	(1,60)	22.64	.0072
Word	(1,60)	118.5	.0092
TaskxSxw	(1,60)	32.01	.0058

are three higher order interactions that must be discussed in order to define the repetition effect in this experiment. The first two must be discussed in order to dispense with them. The third shows us the conditions under which surface form match has an effect on the magnitude of the repetition effect.

The first higher order interaction that involves repetition type is repetition type by task match by target type, which is significant in reaction times ($F(2,120)=4.28$, $MS_e=965$, $p<.05$), as seen in Figure 3, but not in error rates, ($F(2,120)<1$) (means included in Table 2). To analyze the locus of the three way interaction, a simple effects test of the two factor interactions which compose the three way interaction in reaction times (match by repetition type for words and pseudo words) was conducted. This simple effects test showed neither of the two factor interactions to be significant ($F(2,120)=2.34$, $MS_e=965$, $p>.05$ for words, $F(2,120)=2.38$, $MSe= 965$, $p>.05$ for non-words). That is, though the effect of task match for words is significantly different from that of non-words, in neither words nor non-words is the effect of task match on repetition demonstrably greater than chance. Thus, the three factor interaction is caused by two diverging chance variations.

The second interaction of potential interest is found in the three factor interaction of task, target type (word vs non-word) and repetition type (same, different or new). This interaction is significant for both reaction times ($F(2,120)=8.86$, $MS_e=965$, $p<.05$), seen in Figure 4a, and error rates ($F(2,120)=5.25$, $MS_e=.0044$, $p<.05$), seen in Figure 4b. Again, a simple effects test to determine the locus of the interaction was conducted. In this analysis, and all subsequent ones in this set, error rates either showed the same pattern as reaction times or showed no effects, so for clarity of discourse I will discuss only the reaction times. Here there is a significant interaction between target type and repetition type in the lexical decision task($F(2,120)=10.25$, $MS_e=965$),

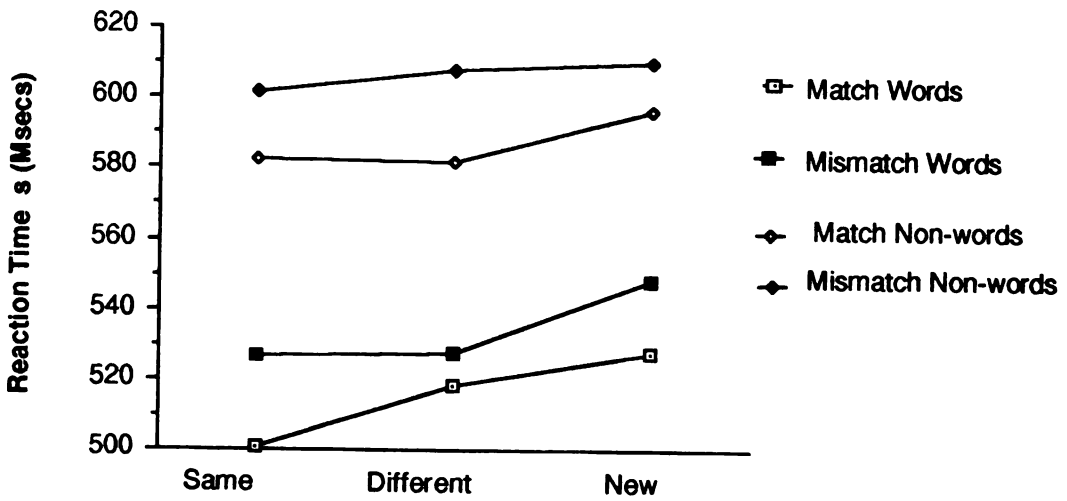


Figure 3: Task Match x Target Type x Repetition

Table 2: Mean Error Rates for Task Match x Repetition x Target Type

	Same Form	Different Form	New Word
Task Match Word	.02	.01	.03
Task Match Non-Word	.12	.11	.16
Task Mismatch Word	.04	.05	.08
Task Mismatch Non-Word	.11	.09	.10

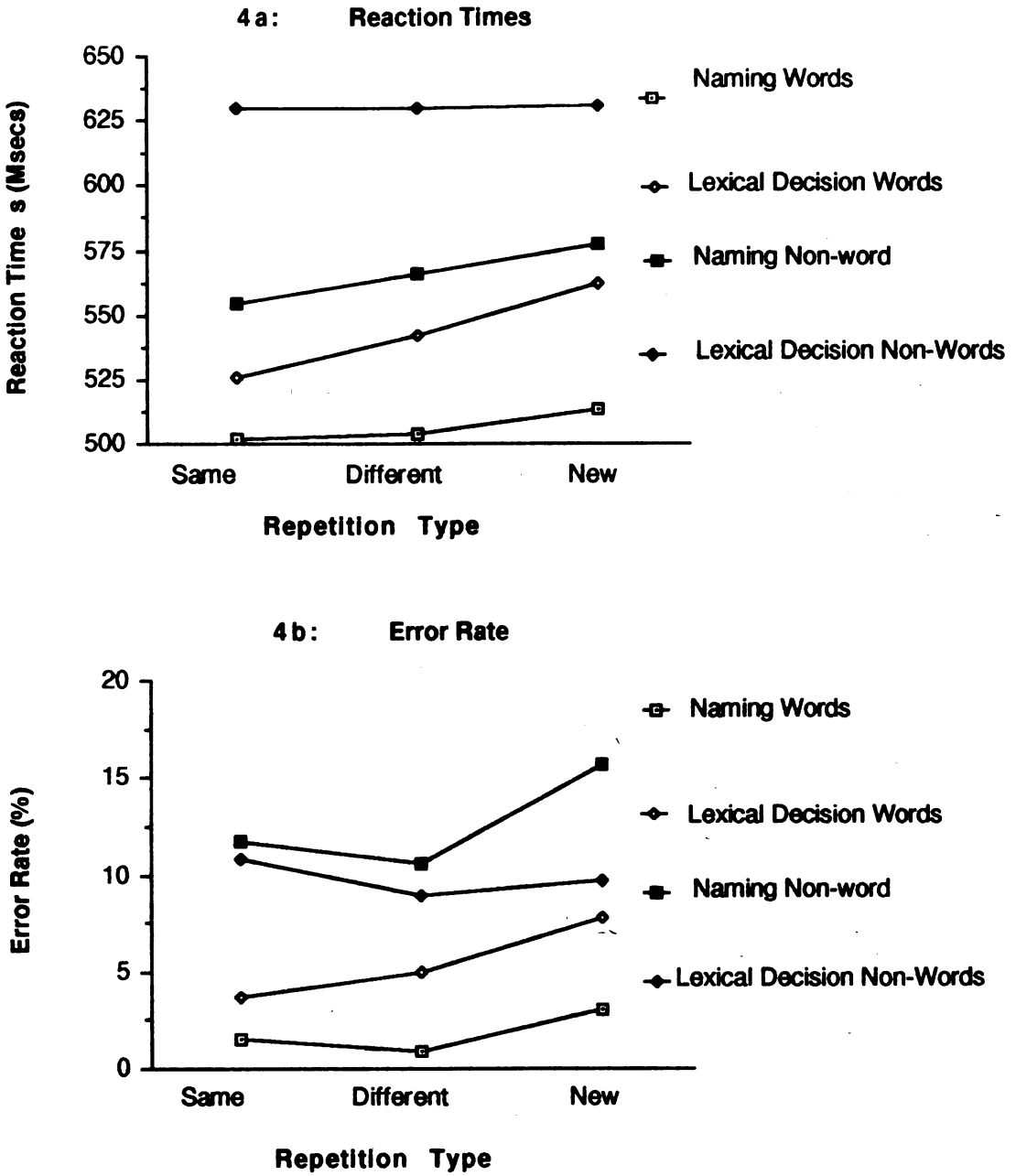


Figure 4: Task x Target Type x Repetition

but not in the naming task ($F(2,120)=1.15$, $MS_e=965$). A second simple effects test of lexical decision data reveals a significant effect of repetition type for words ($F(2,120)=18.34$, $MS_e=1183$, $p<.05$), but not for non-words ($F(2,120)<1$). A comparison of the three levels of repetition type for the naming task and for the words in the lexical decision task using Tukey's H.S.D. test reveals a significant difference between repetitions of the same surface form and new items, but no other significant differences. Thus, it seems that the three factor interaction is the result of the lack of repetition advantage for the non-words in the lexical decision task. This result is similar to that found by Scarborough, Cortese and Scarborough (1977) who found the same pattern of results when 15 items intervene between first and second presentation of a non-word. Because repetitions in a new surface form were statistically indistinguishable from either targets in the old form or new targets, this result is inconclusive with respect to the surface form issues at stake. However, an analysis of the third important interaction will clarify the issue.

The final interaction of importance is the interaction of target type with surface form of target and repetition type seen in Figure 5. This interaction is found only in the reaction time data ($F(2,120)=3.96$, $MS_e=616$, $p<.05$), and for the subsequent simple effects tests, as in the analysis above, only reaction times will be reported. Means for error data are reported in Table 3. It is in the test of the means comprising the repetition effects in this interaction that we find the answers to our questions about surface form effects on repetition. For typed targets, simple effects tests showed the interaction of repetition type and target type was not significant ($F(2,120)<1$). The main effect of repetition for the typed targets produced an $F(2,120)=10.53$ ($MS_e= 926$, $p<.05$). Tukey's HSD test of differences between means in the typed target condition finds the typical abstractionist results, with the repetition effect being uninfluenced by match or

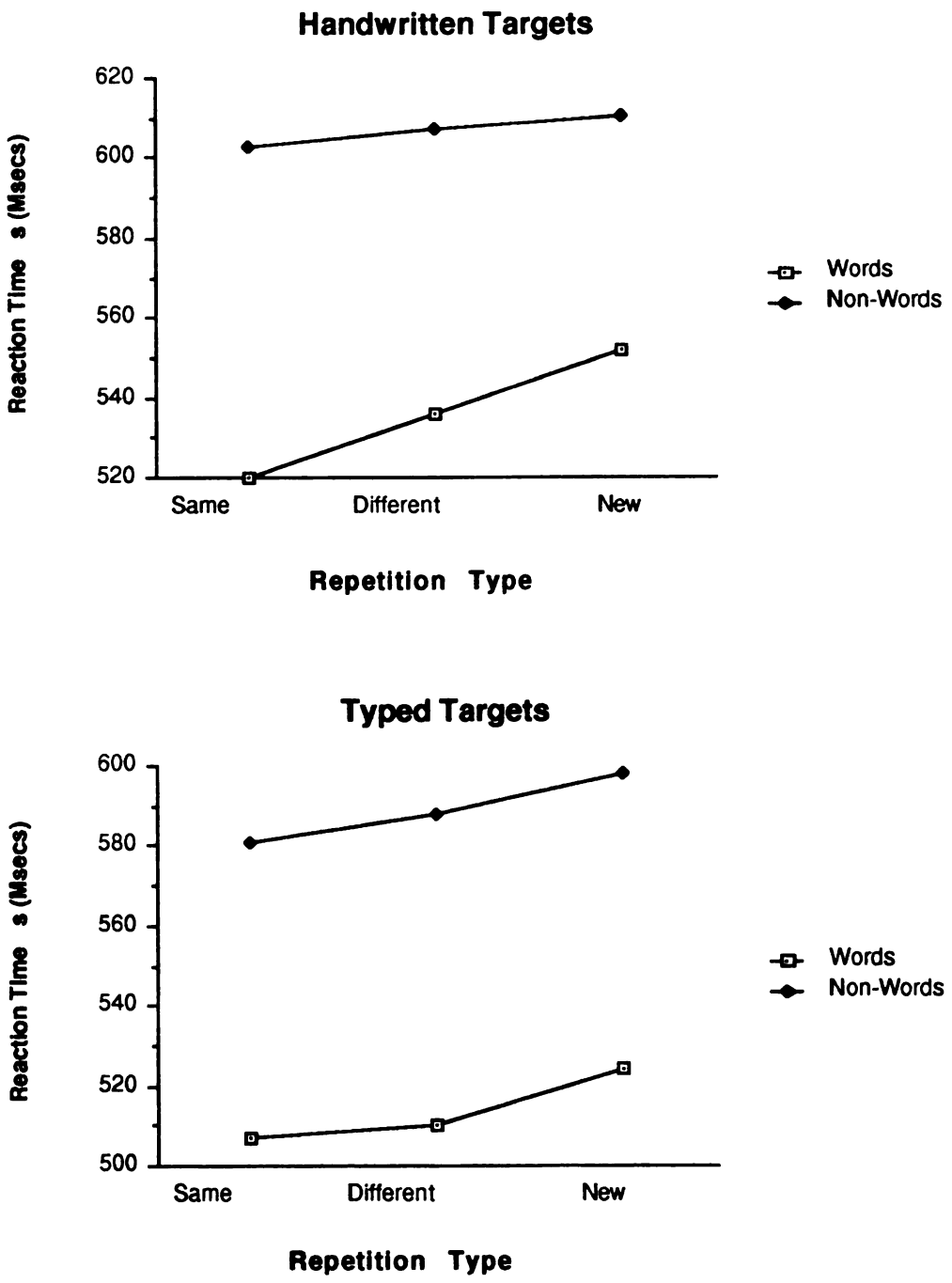


Figure 5: Target Type x Surface form x Repetition

Table 3: Target Type x Surface Form x Repetition

	Same Form	Different Form	New Word
Handwritten Word	.04	.04	.08
Handwritten Non-Word	.13	.10	.15
Typed Word	.02	.02	.03
Typed Non-Word	.09	.09	.10

mismatch of surface form. That is, the means of same and different surface form targets are no different from each other, but both are significantly faster than new targets. In contrast, for handwritten targets, the interaction between repetition type and target type was significant ((F(2,120)=5.21, MS_e=905, p<.05). Separate tests of the repetition effect for words and non-words for the hand written targets found a significant effect of repetition type for words (F(2,120)=18.98, MS_e=836, p<.05) but not for non-words (F(2,120)=1.02, MS_e=836, p<.05). Tukey's H.S.D. test showed that for hand written words, repeated targets whose surface form matched the first exposure were faster than repeated targets of new surface forms, and both were faster than new targets. For handwritten pseudo words, neither type of repeated target was any faster than new targets. Thus, for handwritten targets, repetition effects are isolated in the word targets and the effects found are consistent with instance theories.

An item analysis (F min) was conducted in order to attempt to generalize these results across stimuli as well as subjects. Items with no correct responses in one or more of the experimental conditions were discarded. Because more items were discarded from the non-word list, seperate analyses were conducted for words and non-words. The results of these analyses were generally consistant with the main analysis. For words, a significant interaction of surface form and repetition was found (F(2,152)=3.48, MSe=5135, p<.05). Tukey's H.S.D. test confirmed the patteren found above. Words in the same surface form showed a greater repetition benefit than those in a new surface form, and both were faster than new words. For non-words, the task by repetition was significant (F(2,126)= 4.83, MSe=3899, p<.05) confirming that for non-words, there was no significant effect of repetition in the lexical decision task. No other interactions with repetition were found.

Discussion

The results obtained so far suggest two things. First, the strong form of the task change hypothesis as tested with these tasks appears to be wrong. Changing tasks, at least between naming and lexical decision, does not have any effect on the surface form effects in repetition. In fact, in this experiment, changing tasks had no influence at all on repetition effects, which depend instead on the target task (independent of the first task) and the stimulus surface and lexical characteristics . Second, the effects of surface form match that do occur are not simple.

Not only is the finding that task change has no effect on repetition effects inconsistent with the explanations of surface form match effects that were proposed earlier, but it is also inconsistent with the theories of Kolers discussed in the Introduction. As noted earlier, Kolers predicts that both task change and surface form change should affect repetition effects since the operations performed on the first and second encounter will be different in each case. Here we have a case where Kolers' procedural arguments must predict a reduction in the magnitude of repetition effects. After all, not only are these two tasks logically different from one another and empirically separable, it can also be demonstrated that at least some components of the tasks are neurologically independent.

While it might be tempting to use the lack of task match effects as evidence that the experiment is in some way flawed (through lack of power or statistical error), there are at least two reasons to accept the data as they are found. First, the data are reliable, at least within the same experimental conditions in the same laboratory. To demonstrate this, a secondary analysis of the experiment was run comparing the performance of the first 32 subjects with the second 32. Each group represents a complete replication of the design.

Hence, this analysis constitutes a reliability check of the overall results. The lack of effect of task match on repetition effects was the same in each group (the three way interaction of group, match and repetition produced an $F<1$ in both times and errors). Second, the effect is consistent with results reported by Monsell and Banich (described in Monsell, 1985). In their experiment, equal repetition effects were found when the target task was lexical decision, whether the preceding task was lexical decision or naming. Though the results from this experiment and those of Monsel and Banich disagree with the results found in Logan (1988, 1990), it might be instructive to note that Logan asked subjects to judge the pronounceability of letter strings, rather than actually pronounce those strings. Tentatively then, it seems that this effect is real, and, in its strongest form, the Kolers hypothesis is not true.

Though the task change effects are difficult for most versions of instance theories, the surface form effects found here are a problem for everyone. One might wish to de-emphasize the surface form effects in the pseudowords. The handwritten pseudoword targets showed no repetition effects at all. The typed pseudowords evidenced repetition effects that are statistically consistent with the abstractionist account, but visual inspection indicates that in absolute size, the difference between repetition effects in the same or different surface form are almost as large as the difference between repetition in a new surface form and a completely new target. However, the surface form effects for words are conclusive, and problematic.

For instance theorists, the lack of surface form effects for typewritten word targets is a real problem. This lack of surface form effects for typed stimuli is impossible to explain away as a simple floor effect, since the practice effects are significant. Because instance theories must predict these surface form effects, their absence is serious.

However, just as the absence of surface form effects for typed words is a problem for instance theorists, the presence of such effects for handwritten targets is a problem for abstractionists. There are two things about these data that seem particularly troubling with respect to the findings of Carr et al (1989). Given Carr et al's results, we expected no surface form match effects at all when tasks were matched and in particular when the task was pronunciation. In Carr et al, there were no surface form match effects for words or non-words regardless of the targets' surface form in an oral reading test.

There are several possible explanations for these differences between the present experiment and those presented by Carr et al. Most of these explanations involve the differences between the task used in Carr et al and those used here. Therefore, before discussing the explanations, a quick review of the Carr et al procedure is in order. In the Carr et al procedure, across the various experiments, subjects read aloud texts consisting of about 80 words, random word lists of about the same length laid out on a single page as if they were texts, and texts with familiar content words replaced with pronounceable pseudo words. Each word or non-word on the page came from the same writer and the words or non-words were read in a continuous series. After reading the text or list the first time, subjects immediately read it again. Under these conditions, no surface form effects were found.

One difference between the results found in Carr et al, and those found in Experiment 1 is that Carr et al found their abstractionist results in both words and non-words, while here, repetition effects were inconclusive in typed non-words and non-existent in handwritten non-words. To clarify this discrepancy, a subsidiary analysis was conducted of the condition of the experiment that most closely replicated Carr et al, the condition in which subjects performed the naming task on both exposures. This analysis showed that there were

significant repetition effects in this condition for both words and non-words (Reaction time: $F(2,30)=12.49$, $MS_e= 619$ $p<.05$; Errors: $F(2,30)=4.90$, $MS_e=.0063$, $p<.05$). Further, these repetition effects for non-words followed the same pattern of results as did words in the main analysis. That is, for both words and non-words, the results favored an abstractionist account for typed targets and an instance account for handwritten targets. It appears that the lack of repetition effects in non-words found in the main analysis results from the lack of repetition effects in the lexical decision task for non-words that I reported in the results section. When collapsed across the two target tasks, the practice effects for non-words are washed out.

Another divergence between these results and those of Carr et al is the finding of effects that support the instance account in handwritten targets. One possibility for this lack of convergence with the data collected by Carr et al is discussed in the introduction. The new experiment is considerably more powerful than Carr et al. The surface form effects found in Experiment 1 were quite small, and were found in an experiment that had a great deal of statistical power. The detectable effect size for the Tukey test that found the surface form effect in Experiment 1 was 10.7 msecs, or about 2% of the total reaction time, compared to the 10% effects detectable in Carr et al. Another, and to my mind more interesting, possibility is that the differences between Carr et al and these data are the result of the fact that the surface form of the stimuli was blocked in Carr et al, while it was not blocked in this experiment. That is, the subjects in Carr et al saw a group of words or text created from non-words that were all either handwritten or typed, and if written, they were all written in the same surface form. Subjects in the present experiment were presented with individual words, with handwritten and typed words randomly intermixed. This difference could result in a surface form effect that is not exclusively the result of an

instance influenced processing mechanism, but rather, from two processing mechanisms working in parallel; one a general abstract processor that is restricted to one set of translation parameters at a time and another that relies on a less abstract memory. For instance, in order to obtain repetition benefits for a given string, an abstract system must first convert the string into some sort of standard format (phonological codes, letter identity codes, syllable codes etc). Repetition effects in this system might be unaffected by surface form match or mismatch so long as it can anticipate the orthography of the letter string. However, under conditions in which the system cannot reliably predict the orthography, the convertor might function more slowly or less reliably. Under these conditions, the surface form influenced system might come into play. If this sort of dual processing mechanism is used, then Carr et al may have had surface form match or mismatch effects in the first few words of each set, but these differences were so small as to be undetectable when averaged across the entire block.

There are several reasons that I find this final explanation the most attractive. First, it contains a good explanation of why surface form effects are found for handwritten stimuli but not for typed ones. Remember, the strong form of the instance account has as much problem explaining the lack of surface form effects in typewritten targets as the abstractionist account does explaining the surface form effects in handwritten targets. If there exists an abstract processor of the sort described, it could quite possibly have a default setting to typed scripts. This would not only be generally beneficial for readers who see a great deal of typed script, but in this experiment, where subjects saw two forms of handwriting and only one type face, the system would correctly anticipate the surface form more often if it anticipates a typed text rather than either of the two handwritten forms. If this were the case, then the system would show no surface

form effects when it correctly anticipates the typed target, but fall back on a less abstract surface form dependent system when it encounters an unexpected form. Though Carr et al (1989, Experiment 4) find no effects of subjects expectation about surface form, it could be for the same reason that they did not find surface form effects in transfer to handwritten text, that is, the subjects may have showed increased reading times in the first few items in each paragraph but these effects were lost when the system adapted to the surface form for the rest of the paragraph. An alternative explanation that doesn't require the assumptions about the default values of the system is that the abstract translator is simply incapable of dealing with the new handwriting forms as quickly as it can the typed text with which it is more familiar. This variation of the unusual stimuli argument made in the introduction argues that in Carr et al the system managed to quickly master the new orthographies when presented with an entire set of strings in that orthography, but could not master a set of orthographies it had never seen before and was not allowed to practice sequentially.

In addition to explaining the asymmetry in the surface form effects found in Experiment 1, the explanation offered here is attractive because it obtains some support from outside the literature traditionally cited in the abstractionist/instance debate. Sanocki (1987,1988) found in a series of experiments that subjects can identify briefly presented strings of letters more quickly if the entire string is presented in the same font rather than two different fonts. This suggests that there is some start up time for the letter recognition system to begin translation of a new font. This sort of evidence is consistent with the explanation given above. To test this explanation, Experiment 1 could be replicated blocking the surface form of the target stimuli. If the hypothesis is

correct, match or mismatch of surface form will influence repetition effects for the first string (or several strings) in a block, but not for later strings.

In short, the strongest statements of both the instance theories and the abstractionist theories seem to be wrong. The first predicts effects that do not occur, such as the surface form match/mismatch effects for typed targets; the second fails to predict effects that do occur, such as the match/mismatch effects for handwritten targets. If these data are to be believed, we are required to propose a more complicated theory that allows for the system to produce surface form dependent repetition effects under certain conditions, and surface form independent repetition effects in others. Further, it appears that one such theory, the task change hypothesis, is wrong, at least in its strongest form, as the observed appearance and disappearance of surface form effects is independent of the match or mismatch of the practice and test task.

EXPERIMENT 2

A variation of the first experiment was included in order to both extend the results of Experiment 1, and, further, to test a hypothesis proposed by Kirsner, Dunn, and Standen (1987; Kirsner & Dunn, 1986). Kirsner et al argue that prior surface form will affect later perceptual recognition to the extent that surface form is memorable. Thus, if there is an effect of surface form match on perceptual recognition, then surface form should also be accessible in an explicit memory task. This is perhaps the strongest case for instance theories of repetition benefit. That is, that the repetition benefits not only are by memory traces that are affected by similarity on a number of dimensions normally thought of as "episodic" rather than "semantic", but that these effects are caused by the same traces that result in conscious recall of these details, or (at the very least) these two types of storage are affected by the same variables. Such an elegant explanation seems to demand an empirical test, which is included here.

While it would be tempting to test this sort of theory within subjects, such an undertaking is fraught with peril. Consider that there are two senses in which the Kirsner et al hypothesis could be true. First, items whose surface forms are memorable might be the items that also show surface form effects. Second, conditions which lead to high memorability of surface form would also lead to surface form effects.

In the second case, the basic prediction can easily be tested between subjects. That is, it is the conditions that lead to high memorability that are important. Presumably, those conditions that lead to high memorability for one subject also lead to high memorability for others. In the first case, however, the dangers of testing within subjects outweigh the possible advantages. Consider, if you will, the mechanics of testing this theory within subjects. In whichever

order you administer the two tests, you run the risk of the first test distorting the memory prior to the second test.

If you administer the memory test first, you endanger the comparability of your performance measure to more typical tests of repetition effects by asking subjects to recall the surface form prior to the test. We know that some kinds of failures to recall items, specifically the meaning of newly learned items, can affect performance on related items in tests of perceptual recognition (Dagenbach & Carr , 1990). This being the case, it seems dangerous to add this further complication to an already complicated pattern of repetition effects.

If you administer the memory test after the performance measure, then you are asking subjects to recall the surface forms of items under conditions in which they have seen the items multiple times. The problem with this is that the items are not always appearing in the same surface form. The question of which item the subject is likely to recall is difficult enough, but to then try to decide what this means about the state of their memory at the time of the performance task (which is what you really care about) is simply impossible.

Kirsner et al's theory makes a straightforward prediction about a recall test of surface form taken under the same conditions as those in Experiment 1. Those items or conditions in which surface forms are easily recalled should also show effects of surface form match in the test condition of Experiment 1. Though Kirsner et al (1987) argue that a recall test is the most appropriate explicit memory task, one could also argue in favor of a recognition test. Therefore, both sorts of tests will be included to maximize the generality of the results.

Method

The procedure for Experiment 2 was identical to Experiment 1 with one exception. Instead of two blocks of lexical decision or naming trials, subjects completed the first block of trials followed by one of two types of memorability test. Half of the 64 subjects received a recall test of the surface form of the original practice block. The test was given orally with the list being read aloud and the subject indicating the surface form of each string. The rest of the subjects were given a forced choice recognition test of surface form for each item in which the item was displayed in correct surface form and the incorrect form simultaneously and remained until the subject responded. In both the recall and recognition test, subjects were asked to guess if they could not remember the correct answer. The recognition test has two advantages that argue for its inclusion. First, it is generally more sensitive than recall tests. Second, it is an especially appropriate test for a comparison to repetition studies because it may include ease of perceptual identification (perceptual fluency) as one of its components (Jacoby, 1983a). Thus, if any items are going to show surface form effects, it will be those that are especially memorable in the recognition condition. In both of the test conditions, the test was given without warning, and the instructions in the practice block were identical to those in Experiment 1, thus preventing the use of memory strategies that might lead to different processing in this experiment than that found in Experiment 1.

Results

There are two possible ways to approach the data collected in Experiment 2. The first is to analyze by item, while the other consists of analysis by category. Item analysis would consist of evaluating the memorability of the surface forms of the individual words and non-words. At this point any of a

number of analyses could be undertaken, the most common of which would be to correlate the memorability of each item with the magnitude of the surface form match effect (calculated by subtracting the repetition advantage for trials with mismatched surface forms from the repetition advantage for trials with matched surface forms for each item) found in Experiment 1. If Kirsner, Dunn and Standen are correct, then the magnitude of surface form effect should correlate highly with surface form memorability.

The item analysis proposed above relies on one critical assumption, which must be tested before the correlation can be analyzed. That is, the analysis assumes that there is a substantial correlation of item surface form memorability across subjects. To see this, consider the possibility that the correlation of item memorability between subjects is zero. Under these conditions, subjects from Experiment 1 would have remembered different items than those in Experiment 2, and the correlation of surface form effect and memorability would be zero even if Kirsner et al are correct. In fact, the correlation of surface form effects and surface form memorability would actually consist of the product of the correlation of inter subject memorability and the correlation of surface form and memorability within a given subject. Therefore, correlations were computed for pairs of subjects in each of the four between subject cells of the design (subject task (lexical decision vs naming) and memory test (recognition vs recall)). The correlations were computed using a tetregenous r (r_{tet}) with 158 degrees of freedom (Carroll,1961). Each correlation was conducted between two subjects who had received the same stimulus list. The average correlation in the four cells of the design ranged from .04 to .22 with an average correlation of .08. A list of the individual correlations and the average correlations for each condition are included in Table 4. In

Table 4 (Continued)

Recognition Measures

Lexical Decision Task	Words Only	All Stimuli
P1	.10	-.10
P2	.53	.17
P3	-.58	-.22
P4	.38	.29
P5	.29	.47
P6	-.21	-.31
P7	-.19	.39
P8	-.41	-.19
Mean	-.01	.06
Naming Task		
P1	.11	.02
P2	.07	-.12
P3	.23	-.01
P4	.61	.48
P5	-.17	.29
P6	.35	.44
P7	.21	.29
P8	.45	.34
Mean	.22	.22

Table 4: Inter Subject Correlation Coefficients from Experiment 2

	Recall Measures	
Lexical Decision Task	Words Only	All Stimuli
P1	.41	.32
P2	.04	-.04
P3	.11	.13
P4	-.09	-.07
P5	-.24	-.05
P6	.14	.07
P7	.32	.29
P8	.22	.19
Mean	.11	.10
Naming Task		
P1	.02	.06
P2	-.20	.12
P3	.50	.25
P4	.11	.11
P5	.03	.15
P6	.09	-.01
P7	-.76	-.44
P8	.10	.10
Mean	-.01	.04

addition, a similar correlation was conducted on words alone, with similar results. Correlations ranged from -.01 to .22 with an average correlation of .07.

As I said before, with intersubject correlations this low, it is impossible to derive meaningful correlations between memorability scores collected on one set of subjects and surface form effects collected on another group, since the maximum possible correlation is .08, even if an absolutely perfect relationship exists. Note that other analyses such as a median split of items based on memorability rely on the same basic assumption as the correlational analysis, that is that the more memorable items in Experiment 1 will also be the more memorable items in Experiment 2.

Fortunately, there exists another way to get at the problem that doesn't rely on quite the strong set of assumptions required for the item analysis. While the memorability of individual stimuli is not highly correlated across subjects, there might exist broad categories of items or experimental conditions that lend themselves to higher memorability. For instance, subjects might remember the surface form of words better than non-words. If this were the case, Kirsner, Dunn and Standen's hypothesis predicts that a similar effect should be found in the reaction time data collected in Experiment 1.

In order to test this possibility, a 2 x 2 x 2 ANOVA was conducted on the number of items whose surface form was remembered. The independent variables were type of memory measure (recognition vs recall), target type (words vs non-words) and subject task (naming vs lexical decision). The means for all conditions are included in Table 5. The only significant effect was the memory measure by task interaction ($F(1,60)=6.30$, $MS_e=16.35$, $p<.05$) found in Figure 6. A simple effects test showed a significant effect of task in the recognition measure ($F(1,60)=5.96$, $MS_e=16.35$, $p<.05$) but not in the recall measure ($F(1,60)=1.26$, $MS_e=16.35$,$p>.05$). As seen in the figure, subjects

Table 5: Memory Scores for Experiment 2

	Recall
Lexical Decision	Avg Number Recalled
Words	25.8
Non-Words	23.8
Naming	
Words	26.8
Non-Words	25.1

	Recognition
Lexical Decision	Avg Number Recognised
Words	25.9
Non-Words	25.8
Naming	
Words	23.6
Non-Words	23.2

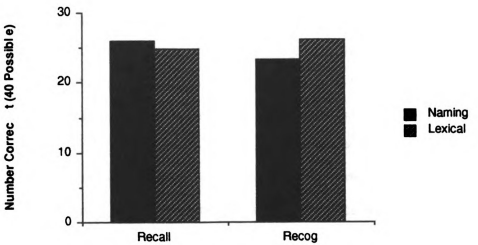


Figure 6: Task x Memory Test

demonstrate a better memory when the initial task performed was a lexical decision task, though memorability scores are generally quite low.

Discussion

In order to interpret the results from Experiment 2, they must be compared to those of Experiment 1. Remember, the prediction of Kirsner, Dunn and Standen (1987) is that the two experiments should show a parallel pattern of results, with surface form effects occurring in those conditions in which surface form was memorable. While the generally low memorability of the surface forms in Experiment 2 would be predicted by very modest surface form effects found in Experiment 1, the other results are more difficult to explain. One especially difficult pattern for the Kirsner et al theory to explain is the results of the recognition memory test. Here, subjects showed greater memory for surface form if they first processed the stimuli in the context of a lexical decision task, rather than in the naming task. The Kirsner et al hypothesis makes the prediction that subjects in Experiment 1 should show the greatest surface form effects when the first task was a lexical decision task. A more conservative hypothesis that still fits the spirit of the Kirsner hypothesis is that these surface form match effects will be found only when both sets of trials used the lexical decision task. To test these predictions, we need to return to the analysis of Experiment 1 and examine the target task by match by repetition interaction. Either of the two versions of the hypothesis require this interaction to be significant, either because the surface form match effects will be found in the naming task mismatch condition and the lexical decision task match condition (for the first formulation of the hypothesis) or in the lexical decision task match condition alone. Either of these patterns would result in a significant three way

interaction. Unfortunately for the Kirsner hypothesis, the relevant interaction is far from significant with an F value of 1.16 (df(2,120), MS_e=925) for the reaction time data and F<1 for the error data.

Caution must be used in evaluating the data, because of the relatively low accuracy of subjects memory for surface form. This low accuracy both reduced the between subjects correlations, thus making the item analysis impossible, and made effects in the ANOVA more difficult to find. However, if the effects of initial task are detectable in memory measures, they must be detectable in reaction times and errors in order for the hypothesis to be true.

GENERAL DISCUSSION

The purpose of the experiments presented here represent and ultimately Quixotic attempt to disentangle the conflicting results in the repetition effects literature. Now that the data are in, what is it that we know about the memory systems responsible for practice effects?

The hypothesis I began with predicted a certain pattern of surface form effects in Experiment 1. It predicted surface form effects when subjects changed tasks, but not when tasks remained the same. Surface form effects did occur in some conditions of the experiment, but, unfortunately, not in the pattern predicted.

Experiment 1 presents new evidence about the conditions under which surface form effects occur. It suggests that the surface form effects are minimal when the target is a standard type font, but small surface form effects can be found with handwritten targets. Two possible explanations can be offered to explain why these effects are found here but not in Carr et al. The less interesting of these explanations involves the higher power of these experiments relative to those of Carr et al. The more interesting explanation relates to the possibility that the surface form effects are found when the system cannot reliably predict the surface form of the stimuli or is not allowed to adapt to that surface form. Tests of this hypothesis are needed before contemplating a complete framework to predict the occurrence of surface form effects. However, taken together, the differences between Experiment 1 and the previous work suggest the correct places to look for the data to begin construction of such a framework. A complete model of repetition effects will need to postulate a system that ignores surface form variation with more typical stimuli or when

allowed to adapt to less typical stimuli, but is effected by surface variation when these conditions are not met.

In addition to pointing out conditions in which surface form effects do occur, Experiment 1 demonstrated some factors that do not change the extent of the system's reliance on surface form. Because the task change hypothesis is clearly not true, at least with these tasks, we are left with the need to discover another organizing principle to make sense of the contradictory repetition effect data. Unfortunately, the alternative explanations are not as straightforward as the task change arguments.

One possibility must be considered before completely abandoning the task change hypothesis. That is, the task change hypothesis may be correct, but only when the tasks vary in with respect to the processes that bring about repetition benefit. One likely candidate for such a shared process that is discussed in the introduction to Experiment 1 is that the lexical decision task and naming task share the same lexical access routes. The fact that repetition benefits transferred completely across the two tasks without respect for surface form in Experiment 1, suggests that this is the locus of the surface form dependent repetition effects, therefore, the task change hypothesis can be saved by a comparatively minor modification. Conceivably, the task change hypothesis is true only when specific tasks are chosen as the target. For instance, the surface form effects found by Roediger and Blaxton occurred as subjects switched from reading to fragment completion and those of Jacoby and Hayman were found in a switch from reading to masked tachistoscopic report. Conceivably, these tasks differ on some more fundamental attribute than do naming and lexical decision. Specifically, theorists such as Norris (1986), Ratcliff and McKoon (1988), or Neely (in press) would lead us to believe that the particular process that lexical decision and naming share are those prior to

lexical access. Therefore, it might be reasonable to require the two tasks to differ in their lexical access routes in order for them to test this revised task change hypothesis.

This interpretation finds some support in the data reported earlier (Monsel, 1985; Ratcliff et al, 1985). It seems that the predictions of Kolers regarding task change are true only when the first and second task are different in certain ways. Therefore, it is not unreasonable to suppose that the same is true of the more elaborate task change hypothesis tested here.

While it is possible to keep the task change hypothesis alive by arguing that a special task change is needed, the lack of theoretical structure to predict which tasks are truly different makes this less than appealing. After all, the lexical decision and naming task seemed quite different at the outset, and it is only when we look at the data that we are tempted to quibble. It seems more reasonable to discard the task change hypothesis and adopt as a working hypothesis the task specific hypothesis from the introduction; the experiments that find surface form effects find them using tasks that are prone to subjects guessing strategies. That is, subjects are more likely to guess that an incompletely recognized word in a perceptual recognition experiment is an old word of the same format as the target than an old word in a different form. This guessing bias could easily explain the surface form effects found in these studies.

The guessing bias hypothesis explains the lack of surface form effects for typed targets in Experiment 1. Consider the tasks used here, naming and lexical decision. Given the low error rates, it is clear that subjects are rarely, if ever, guessing during this experiment. That being the case, we would not expect surface form match effects.

This hypothesis gives some organization to the surface form effect literature. Remember, there remain five clear cases of surface form change effects: Jacoby and Hayman (Experiment 1), Kirsner, Dunn and Standen (1987), Roediger and Blaxton (1987), Jacoby and Witherspoon (1982); and Levy (1983). With the exception of Levy, all of these experiments use tasks subject to guessing bias. The Levy results, while anomalous, are also not replicated in Levy et al (1986).

An additional advantage of this new hypothesis is that it helps explain the pattern of results that gave rise to the Kirsner, Dunn, and Standen hypothesis tested in Experiment 2. If surface form effects are the results of subjects' guessing bias, then it would make sense that the surface form effects would be greatest when subjects remember the surface form. However, when a new test, not prone to guessing bias is used, I found no correlation between subjects memory for surface form and the magnitude of surface form effects. A test of this hypothesis would be possible if an analysis of subjects' incorrect responses were performed on data collected using the masked perceptual identification task. If the guessing hypothesis is correct, then subjects would more often guess items that were previously presented in the same case as the target.

Experiment 2 also offers some insight into the type of memory system responsible for repetition effects. The dissociation of explicit memory for surface form and surface form effects in repetition priming suggests that the two phenomena are supported by different mechanisms, as suggested by Tulving (1983). It seems unlikely that the pattern of results reported by Kirsner, Dunn, and Standen- that is, surface form effects and explicit memory for surface form are correlated- is true for all tasks.

The most important thing that emerges in the course of these experiments is the notion that surface form effects are not a simple, ubiquitous

effect that can easily be used to demonstrate whether repetition effects arise from abstract or instance models of memory. It is becoming increasingly clear that the system itself is not a simple unitary system that uses the same type of representation under all conditions. Neither an entirely instance based nor an entirely abstract representational system can explain the pattern of data presented here, nor can any of the hybrid systems presented elsewhere. Instead, the predictability of the orthographic properties of the stimulus and the typicality of those orthographic properties appear to be operative factors that no current theory emphasizes. As the old cliche goes, more data are needed in order to clarify an increasingly complex picture.

Footnotes

1. Tulving (1985) expresses reservations with classifying these tasks as purely semantic, citing the duration of the priming effects and the independence of priming and recognition among other things. He suggests that priming may be the result of some other memory system that resembles procedural memory. For my purposes though, simply remember that it is not 'episodic' memory we are worried about, but instead, the characteristics of the memory store that supports improvement in performance with practice.

2. Other examples of surface form effects are found in Levy (1983), Kirsner, Dunn & Standen (1987), Roediger & Blaxton (1987), Brooks (1977), Jacoby and Witherspoon (1982) and Kolers (1973,1975,1979).

3. Brooks (1977) finds surface form effects with a script that is not particularly unusual. However, he finds the effect in a visual search task in which subjects are extensively trained on a fairly small set of words (256 in all) in the orthography. This sort of task seems less relevant to subjects normal reading skills.

APPENDIX

Appendix: Stimulus Materials

Words

back 967	should 888	because 883	people 847
state 808	make 794	still 782	long 755
under 707	same 686	might 672	year 660
picture 161	against 626	home 547	found 536
high 497	every 491	left 480	water 442
hand 431	better 414	eyes 401	look 399
point 395	program 394	group 390	toward 386
side 380	order 376	face 371	early 366
case 362	need 360	best 351	power 342
family 331	open 319	problem 313	help 311
human 299	name 294	local 288	today 284
past 281	half 275	field 274	tell 268
money 265	held 264	free 260	special 250
study 246	street 244	south 240	love 232
force 230	voice 226	woman 224	control 223
girl 220	land 217	music 216	child 213
short 212	total 211	plan 205	black 203
play 200	table 19	gone 195	nature 191
fire 187	dark 185	father 183	late 179
hope 178	report 174	heart 173	cold 171

Non word stems

well 897	down 895	those 850	little 831
good 807	world 787	work 760	between 730
never 698	last 676	great 665	house 591
take 611	place 569	small 542	went 507
part 500	school 492	fact 447	think 433
head 424	night 411	find 399	later 397
next 394	city 393	young 385	room 383
present 377	second 373	form 370	white 365
large 361	four 359	church 348	hold 169
country 324	certain 313	door 312	matter 308
above 296	river 165	show 287	feet 283
body 276	week 275	word 274	college 267
rest 163	keep 264	behind 258	office 255
result 244	reason 241	board 239	court 230
wife 228	center 224	common 223	front 221
clear 219	mother 216	party 216	level 213
town 212	class 207	sound 204	hard 202
type 200	road 197	book 193	private 191
ground 186	space 184	return 180	recent 179
brown 176	stage 174	lost 173	tried 170

LIST OF REFERENCES

REFERENCES

Balota, D.A., & Chumbley, J.I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 340-357.

Balota, D.A., & Lorch, R.F. (1986). Depth of automatic spreading activation: Mediated priming effects in pronunciation but not lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 336-345.

Brown, H., Sharma, N. K., & Kirsner, K. (1984). The role of script and phonology in lexical representation. *Quarterly Journal of Experimental Psychology, 36A,* 491-505.

Brooks, L. (1977). Visual pattern in fluent word identification. In Reber, A.S., & Scarborough, D.L. (Eds). *Toward a Psychology of Reading. The Proceedings of the CUNY Conferences* (pp 143-182). Hillsdale, N.J. : Lawrence Erlbaum Associates.

Carr, T. H., Brown, J.S., & Charalambous, A. (1989). Repetition and reading: Perceptual encoding mechanisms are very abstract but not very interactive. *Journal of Experimental Psychology: Learning, Memory, and Cognition. 15,* 763-778.

Carr, T. H., & Pollatsek, A. (1985). Recognizing printed words: A look at current models. In D. Besner, T. G. Waller, & G. E. MacKinnon (Eds.), *Reading*

research: Advances in theory and practice (Vol. 5, pp 1-82). Orlando: Academic Press.

Carr, T. H., Pollatsek, A., & Posner, M. I. (1981). What does the visual system know about words? *Perception & Psychophysics, 29,* 183-190.

Carr T.H., Posner, M.I, Pollatsek, A. & Snyder, C.R.R. (1979). Orthography and familiarity effects in word processing. *Journal of Experimental Psychology: General, 108,* 389-414.

Carroll, J.B. (1961). The nature of the data, or how to choose a correlation coefficient. *Psychometrika, 26,* 347-372.

Carroll, M., & Kirsner, K. (1982). Context and repetition effects in lexical decision and recognition memory. *Journal of Verbal Learning and Verbal Behavior, 21,* 55-69.

Coltheart, M., Besner, D., Jonasson, J.T., & Davelar, E. (1979). Phonological encoding in the lexical decision task. *Quarterly Journal of Experimental Psychology, 31,* 489-507.

Craik, F.I.M., & Kirsner, K. (1974). The effect of speakers voice on word recognition. *Quarterly Journal of Experimental Psychology, 26,* 274-284.

Dagenbach, D., Carr, T.H., & Barnhardt, T. (1990). Inhibitory semantic priming of lexical decisions due to failure to retrieve weakly activated primes.*Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 328-339.

Dixon, P., & Rothkoff, E. Z. (1979). Word repetition, lexical access, and the process of searching words and sentences. *Journal of Verbal Learning and Verbal Behavior, 18,* 629-644.

Jacoby, L. L. (1983a). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning and Verbal Behavior, 22,* 485-508.

Jacoby, L. L. (1983b). Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 21-38.

Jacoby, L. L., & Brooks, L. (1984). Nonanalytic cognition: Memory, perception, and concept learning. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 18, pp 1-47). New York: Academic Press.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110,* 306-340.

Jacoby, L. L., & Hayman, C. A. G. (1987). Specific visual transfer in word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 456-463.

Jacoby, L.L, & Witherspoon, D. (1982).Remembering without awareness. *Canadian Journal of Psychology, 36,* 300-324.

Kirsner,K. (1973). An analysis of the visual component in recognition memory for verbal stimuli. *Memory and Cognition,1,* 449-453.

Kirsner, K., & Dunn, J. (1986). The perceptual record: A common factor in repetition priming and attribute retention. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI* (pp 547-565). Hillsdale, NJ: Erlbaum.

Kirsner, K., Dunn, J., & Standen, P. (1987). Record-based word recognition. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp 147-167). Hove, England: Erlbaum.

Kirsner, K., Milech, D., & Standen, P. (1983). Common and modality specific processes in the mental lexicon. *Memory and Cogniton, 11,* 621-630.

Kirsner, K., & Smith, M. C. (1974). Modality effects in word identification. *Memory & Cognition, 2,* 637-640.

Kolers, P. (1973). Remembering operations. *Memory & Cognition, 3,* 347-355.

Kolers, P. (1975). Specificity of operations in sentence recognition. *Cognitive Psychology, 1,* 289-306.

Kolers, P.A. (1976). Reading a year later. *Journal of Experimental Psychology: Human Learning & Memory, 2,* 554-565.

Kolers, P.A., & Ostry, D.J. (1974). Time course of loss of information regarding pattern analyzing operations, *Journal of Verbal Learning and Verbal Behavior, 13,* 599-612.

Kolers, P., & Roediger, H. L. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior, 23,* 425-449.

Kucera, H., & Francis, W. (1967). *Computational Analysis of Present Day American English.* Providence, R.I.: Brown University Press.

Levy, B.A. (1983). Proofreading familiar text: Constraints on visual processing. *Memory & Cognition,11,* 1-12.

Levy, B. A., & Begin, J. (1984). Proofreading familiar text: Allocating resources to perceptual and conceptual processes. *Memory & Cognition, 12,* 621-632.

Levy, B. A., Newell, S., Snyder, J., & Timmins, K. (1986). Processing changes across reading encounters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 467-478.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review. 95,* 492-527.

Logan, G.D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology, 22,* 1-35.

Masson, M. E. J. (1986). Identification of typographically transformed words: Instance-based skill acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 479-488.

Masson, M. E. J., & Freedman, L. (1985, November). Fluency in the identification of repeated words. Paper presented at the annual meeting of the Psychonomic Society, Boston, MA, USA.

Monsell, S. (1985). Repetition and the lexicon. In A. W. Ellis (Ed.), *Progress in the psychology of language (Vol. 2, pp 147-195).* Hillsdale, NJ: Erlbaum.

Morris, C.D., Bransford, J.D., & Franks, J.J. (1977). Levels of processing versus transfer appropriate processing, *Journal of Verbal Learning and Verbal Behavior, 16,* 529-533.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76,* 165-178.

Morton, J. (1979). Facilitation in word recognition: Experiments causing change in the logogen model. In P. A. Kolers, M. E. Wrolstad, & H. Bouma (Eds.), *Processing visible language* (Vol. 1, pp 259-268). New York: Plenum.

Moscovitch (1981). Multiple dissociations of function in amnesia. In L.S.Cermak (Ed.), *Memory an d Amnesia.* Hillsdale, N.J: Lawrence Erlbaum Associates.

Moscovitch, M., Winocur, G., & McLachlan, D. (1986). Memory as assessed by recognition and reading time in normal and memory-impaired people with Alzheimer's Disease and other neurological disorders. *Journal of Experimental Psychology: General, 115,* 331-347.

Neely, J.H. (in press). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G.

Humphreys (Eds.), *Basic Processes in Reading: Visual Word Recognition*. Hillsdale, N.J: Lawrence Erlbaum Associates.

Norris, D. (1986). Word recognition: Context effects without priming, *Cognition, 22*, 93-136.

Patterson, K.E., & Marcell, A.J. (1977). Aphasia, dyslexia, and the phonological coding of written words. *Quarterly Journal of Experimental Psychology, 29,* 307-318.

Posner, M.I., & Keele, S.W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology,77,* 353-363.

Posner,M.I., Goldsmith,R., & Welton, K.E. (1967). Perceived distance and the classification of distorted patterns. *Journal of Experimental Psychology,73,* 28-38.

Ratcliff, R., Hockley, W., & McKoon, G. (1985). Components of activation: Repetition and priming effects in lexical decision and recognition. *Journal of Experimental Psychology: General, 114,* 435-450.

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review, 85*, 385-408.

Ratcliff, R., McKoon, G., & Verwoerd, M. (1989). A bias interpretation of facilitation in perceptual identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 378-387.

Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across fixations. *Cognitive Psychology, 12,* 206-226.

Roediger, H. L., & Blaxton, T. A. (1987). Effects of varying modality, surface features, and retention interval on priming in word-fragment completion. *Memory & Cognition, 15,* 379-388.

Sanoki, T. (1987). Visual knowledge underlying letter perception: Font-specific, schematic tuning. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 267-278.

Sanoki, T. (1988). Font regularity constraints on letter recognition. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 472-480.

Scarborough, D. L., Cortese, C., & Scarborough, H. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 1-17.

Scarborough, D. L., Gerard, L. & Cortese, C. (1979). Accessing lexical memory: The transfer of word repetition effects across task and modality. *Memory & Cognition, 7,* 3-12.

Schvaneveldt, R.W., & McDonald, J.E. (1981). Semantic context and the encoding of words-evidence for two models of stimulus analysis.*Journal of Experimental Psychology: Human Perception and Performance,7,* 673-687.

Seidenburg, M.S., Waters, G.S., Sanders, M., & Langer, P. (1984). Pre- and postlexical loci of contextual effects on word recognition. *Memory and Cognition, 12,* 315-328.

Seidenberg, M. S., & McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 98,* 523-568.

Smith, E.E., & Spoehr, K.T. (1974). The perception of printed English: A theoretical perspective. In B.H. Kantowitz(Ed.), *Human Information Processing: Tutorials in Performance and Cognition.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Srinivas, K., & Roediger, H.L. (1990). Classifying implicit memory tests: Category association and anagram solution. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 389-412.

Thompson, D.M., & Tulving, E. (1970). Associative encoding: Weak and strong cues. *Journal of Experimental Psychology*, *86*, 255-262.

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

Tulving, E. (1983). *Elements of episodic memory.* New York: Oxford press.

Tulving, E. (1985). How many memory systems are there? *American Psychologist*, *40*, 385-598.

Tulving, E., & Thompson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352-373.

West, R.F. & Stanovich, K.F. (1982). Source of inhibition in experiments on the effect of sentence context on word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*, 385-399.