

4-14-2010

## The Effects of Physical Distinctiveness and Word Commonness on Brain Waves and Subsequent Memory: An ERP Study

Siri-Maria Kamp  
*University of South Florida*

Follow this and additional works at: <https://digitalcommons.usf.edu/etd>



Part of the [American Studies Commons](#)

---

### Scholar Commons Citation

Kamp, Siri-Maria, "The Effects of Physical Distinctiveness and Word Commonness on Brain Waves and Subsequent Memory: An ERP Study" (2010). *USF Tampa Graduate Theses and Dissertations*.  
<https://digitalcommons.usf.edu/etd/1675>

This Thesis is brought to you for free and open access by the USF Graduate Theses and Dissertations at Digital Commons @ University of South Florida. It has been accepted for inclusion in USF Tampa Graduate Theses and Dissertations by an authorized administrator of Digital Commons @ University of South Florida. For more information, please contact [digitalcommons@usf.edu](mailto:digitalcommons@usf.edu).

The Effects of Physical Distinctiveness and Word Commonness on Brain Waves and Subsequent  
Memory: An ERP Study

by

Siri-Maria Kamp

A thesis submitted in partial fulfillment  
of the requirements for the degree of  
Master of Arts  
Department of Psychology  
College of Arts and Sciences  
University of South Florida

Major Professor: Emanuel Donchin, Ph.D.  
Geoffrey F. Potts, Ph.D.  
Kenneth J. Malmberg, Ph.D.

Date of Approval:  
April 14, 2010

Keywords: Word Frequency, Event-Related Potential, P300, Frontal Slow Wave, Recall Memory

© Copyright 2010, Siri-Maria Kamp

## **Acknowledgments**

I would like to thank my advisor Emanuel Donchin for all of his support throughout this study. Michelle-Chin Quee deserves acknowledgement for helping with data collection and entry. Furthermore, I would like to thank the other members of the Cognitive Psychophysiology laboratory, especially Ty Brumback, Geoffrey Potts and Yael Arbel, as well as Kenneth Malmberg and his lab, for their helpful comments and suggestions in lab meeting discussions.

## Table of Contents

List of Figures . . . . .	iii
Abstract . . . . .	iv
Chapter 1 Introduction . . . . .	1
1.1 Effects of Distinctiveness on Memory . . . . .	1
1.2 Effects of Word Frequency on Memory - Another Type of Distinctiveness? . . . . .	5
1.2.1 Operational Definition of Word Frequency . . . . .	6
1.2.2 Effects of Word Frequency on Cognition and Memory . . . . .	7
1.2.3 Are LF Words Distinctive? . . . . .	8
1.2.4 Do HF Words Enhance Item-to-Item Elaborative Processes? . . . . .	10
1.3 The P300, the Frontal Slow Wave, and the N400 . . . . .	11
1.3.1 The P300 as a Measure of Distinctiveness . . . . .	12
1.3.2 Slow Waves as Index for Working Memory Processes . . . . .	13
1.3.3 The N400 as an Index of Semantic Integration . . . . .	15
1.4 Correlations of ERP Amplitudes and Memory for Distinctive Stimuli . . . . .	15
1.5 ERP- and Other Neuroimaging Studies on Word Frequency Effects . . . . .	20
1.6 Objectives and Hypotheses for the Current Study . . . . .	24
Chapter 2 Methods . . . . .	27
2.1 Participants . . . . .	27
2.1.1 Incentives for Good Performance . . . . .	27
2.2 Materials . . . . .	28
2.2.1 Stimuli . . . . .	28
2.2.2 Word Lists . . . . .	29
2.2.3 Recall Sheets . . . . .	30
2.2.4 Individual Word Familiarity Questionnaire . . . . .	30
2.2.5 Stimuli and Design of the Oddball Task . . . . .	30
2.3 Task and Procedure . . . . .	31
2.4 EEG Recording and Analysis . . . . .	32
2.5 Statistical Analysis . . . . .	33
Chapter 3 Results . . . . .	35
3.1 Debriefing . . . . .	35
3.2 Behavioral Data . . . . .	35
3.2.1 Recall Rates . . . . .	35
3.2.2 Analysis of the Subjective Familiarity Questionnaire . . . . .	36

3.2.3	Analysis of Output Order . . . . .	37
3.3	Event-Related Potentials . . . . .	38
3.3.1	PCA Results . . . . .	42
3.3.1.1	Frontal Factor . . . . .	43
3.3.1.2	Right Lateralized Spatial Factor . . . . .	45
3.3.1.3	Left Lateralized Spatial Factor . . . . .	45
3.3.1.4	Parietal Spatial Factor . . . . .	48
3.4	Oddball Task . . . . .	50
Chapter 4	Discussion . . . . .	54
4.1	Summary of Results . . . . .	54
4.2	Correlations of ERP Amplitudes and Recall in the Von Restorff Paradigm . . . . .	56
4.3	Word Frequency Effects and Their Reversal due to List Composition . . . . .	59
4.4	Common Processes of the Von Restorff Effect and Word Frequency Effects? . . . . .	62
4.4.1	ERP Components and Correlations to Recall for LF Isolates . . . . .	63
4.4.1.1	Comparison of Output Order for LF Isolates and Size Isolates . . . . .	66
4.4.2	ERP Components and Correlations to Recall for HF and LF Standards . . . . .	67
4.4.3	Correlations of N400 Amplitudes and Recall Success for HF Isolates . . . . .	68
4.4.4	Subsequent Memory Effects Across Word Types . . . . .	70
4.4.5	Implications of our Data for Word Frequency Effects on Recall . . . . .	70
4.5	Implications of the Current Study for the Frontal Slow Wave . . . . .	72
4.6	Abnormalities in the Current Study's Oddball Paradigm . . . . .	74
Chapter 5	Conclusions and Future Directions . . . . .	76
References	. . . . .	78

## List of Figures

Figure 1	Serial position curves. . . . .	36
Figure 2	Subjective familiarity ratings. . . . .	37
Figure 3	Output order analysis. . . . .	38
Figure 4	Grand averages for Von Restorff lists. . . . .	39
Figure 5	Grand averages for HF lists including LF isolates. . . . .	40
Figure 6	Grand averages for LF lists including HF isolates. . . . .	41
Figure 7	Spatial Factor Loadings. . . . .	42
Figure 8	Spatial factor 1. . . . .	44
Figure 9	Spatial factor 2. . . . .	46
Figure 10	Spatial factor 3. . . . .	47
Figure 11	Spatial factor 5. . . . .	49
Figure 12	Grand average ERPs from the oddball task. . . . .	51
Figure 13	The first three spatial factors from the oddball task. . . . .	53

**The Effects of Physical Distinctiveness and Word Commonness on Brain Waves and Subsequent Memory: An ERP Study**

**Siri-Maria Kamp**

**ABSTRACT**

Words that deviate in their physical characteristics from their surrounding lead to enhanced recall memory, a pattern known as the *Von Restorff effect*. Furthermore, common (high frequency; HF) words are more likely to be recalled than uncommon (low frequency; LF) words when they occur in pure lists, while this pattern is reversed in mixed lists of both HF and LF words. This study investigated whether the Von Restorff effect and the reversal of word frequency effects in mixed lists, which may both be explained by enhanced perceived *distinctiveness*, are associated with common underlying brain processes. Event-related potentials (ERPs) were recorded while participants studied and subsequently recalled 70 word lists using rote memorization strategies. The three list types included (1) 14 regular-sized and one larger word, (2) 14 HF words and one LF word, or (3) 14 LF words and one HF word. The behavioral data showed a typical Von Restorff effect, a word frequency effect, as well as a reversal of the word frequency effect for LF words isolated in HF word lists (“LF isolates”). Larger words and LF isolates elicited a *P300*, an ERP component associated with subjective distinctiveness, whose amplitude was correlated with subsequent recall for both word types. This indicates that LF isolates were perceived as distinctive, and that this perceived distinctiveness aided subsequent recall in a similar way as for physically deviant words. Both larger words and LF isolates also elicited a left-lateralized slow wave which was larger for subsequently recalled than for not recalled words. This ERP component supposedly reflects item-to-item elaborative processes, indicating that such elaborative processes are enhanced when LF words occur in HF word list. HF words isolated in lists of LF words did not elicit comparable ERP subsequent memory effects. Rather, for these “HF isolates”, the *N400* was negatively correlated with subsequent recall,

an ERP component that reflects semantic integration processes. We conclude that the reversal of the word frequency effect in mixed lists can be explained by a combination of enhanced subjective distinctiveness and enhanced inter-item elaborative processes for LF words that occur in lists of HF words.



## Chapter 1

### Introduction

The goal of this study is to investigate similarities and differences between the effects of two types of “distinctiveness” on memory. First, it is well known that items that are physically deviant from their surrounding are more likely to be correctly recalled on a later test. Furthermore, how commonly a word occurs in a language also has very specific effects on both recall and recognition memory. Both of these effects on recall memory can be explained in terms of distinctiveness theories. The current study compares the neuro-cognitive processes associated with these two types of distinctiveness effects to determine whether common processes underlie the two effects on recall, or whether the term “distinctiveness” is used for entirely different processes in the two cases.

We measure brain processes by the means of event-related potentials (ERPs) and attempt to determine the conditions under which they correlate with performance on subsequent recall memory tests. We focus on two ERP components known as the *P300* and the *frontal slow wave*. Previous studies indicate that the amplitudes of these two components are associated with subsequent recall under different conditions – the *P300* when items are physically distinctive and when participants do not use elaborative memorization strategies; and the frontal slow wave when they do use elaborative strategies and for non-distinctive items. The current study attempts to clarify which of the two components is correlated with subsequent recall when words exhibit different levels of commonness; that is, for *word frequency effects* on recall memory.

#### 1.1 Effects of Distinctiveness on Memory

The question why humans remember some facts and experiences, but not others has puzzled many researchers for a long time. One repeatedly reported finding is that items that differ strongly from their temporal or spatial neighborhood – in other words, items that are *distinctive* – exhibit an advantage in later memory tests. The German Gestalt psychologist Hedwig Von Restorff (1933)

systematically manipulated distinctiveness in a list learning paradigm and showed that, for example, a number that is studied in a sequence of nonsense syllables is better remembered than the syllables that are studied in the same sequence. The same memory enhancing effect is observed for isolated syllables that are studied in a sequence of numbers, as well as for lists composed in a similar way of other types of material such as figures, colored squares and letters (Von Restorff, 1933). This effect, which was later named the *Von Restorff effect* (or isolation effect), is usually strong when memory is retrieved in free recall tests, but is often reported to be absent or weak for standard recognition tests (McLaughlin, 1968; Von Restorff, 1933). It is worth noting that Von Restorff's primary goal was not to study memory *per se*. Rather, she interpreted her findings in terms of figure-ground separation ideas originating in the field of Gestalt psychology. However, an important question following from the finding of enhanced recall for distinctive stimuli is *why* human memory is better for events that "stand out" from their environment.

Theories differ in their claims on the time point at which distinctiveness affects memory. That is, distinctive items may be differentially well encoded at study time, or alternatively, distinctiveness of study items may affect retrieval processes at the time of the memory test (McDaniel & Geraci, 2006). The hypothesis that the effects of distinctiveness operate at the time of retrieval are in part based on the finding that memory for distinctive items is enhanced when one of the first items in the list is the distinctive one. At this time, the individual cannot yet notice the item's distinctiveness (Von Restorff, 1933), leading to the conclusion that the processes leading to enhanced memory occur at a later time. In contrast, one widely accepted and applied notion is that distinctiveness operates at study time because physically distinctive items attract more attention than non-distinctive items, which leads to better encoding and therefore the memory advantage (McDaniel & Geraci, 2006; Schmidt, 1991). As will be explained later in this chapter, strong evidence that, at least in part, distinctiveness effects on memory operate already at the time of study comes from event-related potential (ERP) studies. Thus, several previous ERP studies indicate that the amplitudes of a distinctiveness-related ERP component, the so-called *P300* elicited at study time, are correlated to subsequent recall success under certain conditions. It is worth noting that the mechanisms suggested by different theories are not necessarily mutually exclusive. Therefore, although the focus of the current study is on processes operating at the time of encoding, it seems likely that specific processes at retrieval time also provide a contribution to distinctiveness effects on memory.

In order to explain effects of distinctiveness on memory, a relevant distinction is between *item-specific* and *relational processing* (Hunt & Einstein, 1981; Humphreys, 1976). Item specific processing refers to the encoding of features that are unique to one study item in a group. These may be physical or semantic features that can be detected at the time the stimulus is encountered. Relational processing, in turn, involves relating the study item to other information in memory. This may include forming associations between study items or connecting a study item to the context or information retrieved from long term memory. Hunt and Einstein (1981) reported that the two types of processing elicited at study time make separable contributions to successful memory. According to their findings, whether intensive in item-specific or relational processing is most beneficial for subsequent memory depends on the nature of the items in the study list, such that unrelated study lists benefit more from relational processing instructions than semantically interrelated study lists. Furthermore, the Hunt and Einstein report that optimal memory performance is achieved when both item-specific *and* relational processing are engaged at time of encoding.

Of interest for the current discussion is the notion that the cognitive processing of a distinctive item involves enhanced item-specific processing, because this item exhibits unique features that distinguish it from the other items in the group in which they are studied. That is, if an item is distinctive compared to its neighborhood, more processing is directed towards these unique features. This, in turn, leads to a distinctive memory trace that is well distinguishable from other memory traces. However, an item can only “stand out” from its study group if the rest of the study items are processed as similar to each other. For example, a syllable in a list of numbers will “stand out” and therefore later be recalled more easily, but a syllable in a heterogeneous list of a number, a figure, a letter and a color will not (Von Restorff, 1933). Therefore, distinctiveness can be defined as “processing difference in the context of similarity”, which implies that both strong relational processing among all surrounding, similar items as well as enhanced item-specific processing of the distinctive item takes place (Hunt, 2006). An important implication of this argument is that the concept of distinctiveness refers to a psychological process – the subjective processing of difference in the context of similarity – rather than to a characteristic that is intrinsic to the study item (Hunt, 2006; Donchin & Fabiani, 1991). At this point it is worth noting that the so-called *oddball paradigm*, which is used to study one brainwave component of interest for the current study, the *P300*, can be thought of as manipulating distinctiveness according to the definition given here. Thus, this paradigm contains in-

frequent, “distinctive” events on the background of frequent, similar events. The oddball paradigm and the suggestion that the P300 can hence be considered a measure of subjective distinctiveness will be discussed in detail later in this chapter.

Hunt’s definition of distinctiveness seems intuitive – if an item is distinctive, it must be distinctive relative to something else. For the current study it is, however, important to note that distinctiveness is not always defined on the basis of the immediate context of a study item, but can be defined relative to previous experiences. One example of this latter type of distinctiveness are words that occur very uncommonly in a language. These words can be thought of as “distinctive” compared to other words in lexical memory, because they have been previously encountered less frequently, because they have less common features or fewer connections with other items in memory (a more detailed discussion of what makes uncommon words distinctive follows later in this chapter). That is, distinctiveness in the case of uncommon words is defined on the basis of fewer previous experiences with these words compared to other, more common words, rather than based on the immediate context. It is this type of distinctiveness that has been named *secondary distinctiveness* (Schmidt, 1991). A relevant question is whether words of different frequencies behave similar in an *oddball paradigm* compared to words that have deviant physical properties. This question will be discussed later in this chapter.

The example of using the term *distinctiveness* to describe words uncommonly encountered in a language demonstrates that “distinctiveness” is not a unitary attribute. Different types of distinctiveness have been identified. Thus, Schmidt (1991) distinguishes between *emotional distinctiveness*, *processing distinctiveness*, *primary distinctiveness* and *secondary distinctiveness*. As noted before, of interest for the current study are primary and secondary distinctiveness. Schmidt defines primary distinctiveness as distinctiveness with respect to the immediate context. Effects of primary distinctiveness can only be observed in within subject manipulations, when “non-distinctive” items are present. A typical example of effects of primary distinctiveness on memory is the Von Restorff effect discussed above. Secondary distinctiveness, in turn, is defined with respect to the pool of all prior experiences, and one instance are words that occur uncommonly in a language, as described above. According to Schmidt, secondary distinctiveness also comprises the “bizarreness effect”, which refers to the phenomenon that under certain conditions, bizarre study items are better recalled than plausible ones (Worthen, 2006; Davidson, 2006). Further examples of secondary

distinctiveness are irregularly spelled words or pictures of unusual faces which are often better remembered than their counterparts (Schmidt, 1991). An open question is whether the different types of distinctiveness are related instances of the same phenomenon, or whether the same term is used to describe entirely different processes that are related only on a conceptual level. One way of approaching this issue is by investigating and comparing the brain processes associated with both types of phenomena. This idea is reflected in the research question of the current study: whether the effects of secondary distinctiveness on recall memory, as manipulated by word commonness, are associated under some conditions with similar neuro-cognitive processes as the effects of primary distinctiveness on memory. Alternatively, the two memory phenomena could be based on entirely dissociable processes.

It is clear from this line of thought that a brain wave signature that captures the effects of primary distinctiveness on memory, as observed in the Von Restorff effect, could help clarify our question. Thus, if this brain wave was also related to memory processes for secondary distinctiveness effects, one could conclude that the two types of distinctiveness are subserved by shared brain processes. One brainwave that has been implicated as a measure of subjective distinctiveness - the *P300* - will be the main focus of this study. A discussion of the characteristics and functional significance of this component, as well as its relationship to memory processes, follows after a more detailed review of theories and findings on word frequency effects.

## **1.2 Effects of Word Frequency on Memory - Another Type of Distinctiveness?**

How commonly a word occurs in a language, or how frequently it is used - in short its *word frequency* - is a property of linguistic material that is often studied in its effects on cognitive processes. When studying the effects of word frequency on cognition and memory, it is first necessary to provide an operational definition of the construct. Different possibilities for accomplishing this are discussed in the next subsection. Afterwards, we will review the effects of word frequency on cognition and memory as well as provide two groups of possible explanations for these effects. These two groups of possible explanations form the hypotheses in the current study.

### 1.2.1 Operational Definition of Word Frequency

The construct of interest for many studies on word frequency effects is how frequently an item has been previously encountered. There are different kinds of stimulus material that can be used to operationalize this construct. Probably the most commonly used stimulus set on word frequency has been developed by Francis and Kucera (1982). These authors counted the number of occurrences of a large number of words in a big sample of written material (Francis & Kucera, 1982). Using these frequency norms, one can operationally define LF words as words that occur less than a certain number of times in a million. HF words are defined analogously as words that occur more than some number in a million. While Francis and Kucera's frequency norms are a very commonly used means of measuring word frequency, it is worth noting that they have been criticized for a potential sampling-bias due to the inclusion of only written material (Balota, Pilotti, & Cortese, 2001).

Other researchers have relied on word familiarity norms to operationalize participant's previous exposure to words. Probably the most commonly used norms for this purpose were developed by Toglia and Battig (Toglia & Battig, 1987). These authors had a large number of participants rate their perceived familiarity of words (among other characteristics) on a scale from 1 (*very unfamiliar*) to 7 (*very familiar*). Some authors suggest that word familiarity is a more accurate measure of frequency of usage represented in the mental lexicon than word frequency norms based on only written material (Connine, Mullennix, Shernoff, & Yelen, 1990).

There are advantages and disadvantages with either one of these measures of prior exposure. As noted above, Francis and Kucera's norms consider only written material. The familiarity ratings by Toglia and Battig, in turn, may be influenced by semantics, imageability and other properties of the words because participants may rely on these features to perform the ratings (Balota et al., 2001). Because pilot studies suggested that both familiarity and word frequency norms play an important role, the choice of stimulus material for the current study is as combination of subjective familiarity ratings (Toglia & Battig, 1987) and word frequency norms (Francis & Kucera, 1982).

One complication with using any of the norms to select experimental stimuli is between subject variability of previous exposure. It is readily apparent that no norm precisely reflects every subject's prior exposure to each individual word. To capture these individual differences, the current study measures the degree of familiarity of a large subset of the experimental stimuli by administering a subjective familiarity questionnaire at the end of the experiment.

### 1.2.2 Effects of Word Frequency on Cognition and Memory

A well-replicated observation is that word frequency affects the speed of a word's recognition as a word (vs. a pronounceable non-word) in lexical decision tasks (Mason, 1976; Polich & Donchin, 1988; Forster & Chambers, 1973) as well as latencies in word naming tasks (Forster & Chambers, 1973; Gerhand & Barry, 1999). This difference in processing speed dependent on word frequency is commonly attributed to slower lexical access of uncommon, or low frequency (LF) words compared to common, or high frequency (HF) words (Forster & Chambers, 1973; Connine et al., 1990).

Beyond this effect of word frequency on processing speed, very specific influences of word frequency on memory have been reported. The precise effects of word frequency on memory depend on the type of test used to assess memory. Probably most commonly studied is the word frequency effect on recognition memory, referring to the well-replicated finding that LF words are recognized more accurately in subsequent memory tests than HF words. More specifically, during recognition memory tests, participants are more likely to correctly judge LF words as "old" (previously studied), while at the same time producing fewer false alarms (judgements of "old" to previously not studied words) for LF words, compared to HF words (Mandler, Goodman, & Wilkes-Gibbs, 1982). This pattern is known as the *mirror effect* of word frequency on recognition memory (Glanzer & Adams, 1990, 1985; Malmberg & Nelson, 2003). Qualitatively, the word frequency effect on recognition is usually obtained regardless of the proportion of HF and LF words in a study list, but study list composition does appear to affect the *magnitude* of the differences in recognition accuracy between LF and HF words (Malmberg & Murnane, 2002).

For free- and serial recall tests, reports of word frequency effects are less consistent across studies. Most researchers report that when studied in "pure" or "blocked" lists containing only HF or only LF words, HF words lead to higher recall levels than LF words (Mandler et al., 1982). In mixed lists including both HF and LF words, however, the two word types are recalled equally well (Watkins, LeCompte, & Kim, 2000; Fernández, Klaver, Fell, Grunwald, & Elger, 2002; Hulme, Stuart, Brown, & Morin, 2003). Other researchers report that in mixed lists, uncommon words are more likely to be recalled, while this difference disappears in blocked lists (McDaniel & Geraci, 2006). Although the data from these different studies disagree on the specific pattern of the effect of word frequency on recall memory, they are consistent in that LF words in mixed lists are always better remembered than LF words in blocked lists, while the reverse is true for HF words. The consistency of this

finding across studies indicates that the differential effects of word frequency on recall memory depend strongly on the immediate context at study time. The cognitive processes leading to this abolition, or reversal, of the HF advantage on recall memory when mixed lists are studied, are the main focus of this study.

Thus, the question of interest for the current study is which cognitive processes are qualitatively or quantitatively different between mixed and pure lists, leading to the abolition or reversal of the word frequency effect on recall. There are different possible explanations for this reversal effect, which can be classified into two broad groups for the purpose of this study. The first class of explanations states that LF words are *distinctive* compared to HF words. The subjective perception of their distinctiveness within the list is enhanced in lists where HF words (that is, less distinctive items) are present. According to this idea it is the enhancement in perceived distinctiveness that, similar to the Von Restorff effect, then facilitates memory retrieval for LF words embedded in HF word lists. Although the two proposals are not mutually exclusive, the alternative group of theories poses that the presence of HF words in a list facilitates elaborative or organizational processes. According to this idea, each word, regardless of its frequency, equally benefits from the presence of other HF words in the list because these provide more opportunities for item-to-item elaboration. The outcomes of these elaborative and organizational processes can then be used to guide and facilitate memory retrieval during the recall test. Therefore, both of these groups of explanations can account for the finding that HF words are recalled better in lists of only HF words, while LF words are recalled better in mixed lists, where many HF words are present.

A more detailed review of the two groups of explanations for the word frequency effect on recall and its reversal for mixed study lists follows next.

### **1.2.3 Are LF Words Distinctive?**

The idea that LF words are “distinctive” compared to HF words is widely accepted and most often used to explain word frequency effects on recognition memory within different theoretical frameworks. The definition of this type of “distinctiveness” is, however, not necessarily identical to what we refer to when discussing the Von Restorff effect and what we suggest can be measured by the P300 (see next chapter). Therefore, to avoid confusion, this section will refer to LF words as more *unique*, rather than *distinctive*, compared to HF words. In fact, whether or not the uniqueness



that is often attributed to LF words and the distinctiveness of items that physically stand out from their context share underlying neuro-cognitive processes that correlate to memory under certain conditions is the main research question of the current study.

The obvious reason why LF words could be considered unique is that people do not frequently encounter them in their daily life. Thus, LF words have been previously experienced less often and therefore are likely to be less familiar than HF words. Besides this readily apparent explanation of LF word's uniqueness, there are a number of other suggestions for characteristics that co-vary with word frequency and that cause LF words to be more unique. Thus, the frequency of prior exposure to a word may not be the only explanation for LF word's uniqueness and thereby for the effects of word frequency on cognitive processes such as recognition and recall memory.

A different possibility is that LF words tend to have unusual features that are not shared by many other items in memory. Thus, they are more likely to exhibit irregular spelling, hence containing more unusual letter combinations (Criss & Malmberg, 2008), and more unusual phonemes (Landauer & Streeter, 1973) than HF words. A different possible source for a LF word's uniqueness is that there is a tendency that LF words have been previously encountered in fewer different contexts than HF words (i.e., they exhibit lower *context variability*) (Steywers & Malmberg, 2003; Adelman, Brown, & Quesada, 2006).

Yet another possibility is that LF words attract more attention, or require more attentional processing resources to be encoded episodically. Several theories explicitly incorporate that increased attentional resources are allocated to item-specific information of LF words compared to HF words, which may be an important influence on word frequency effects (Glanzer & Adams, 1990; Malmberg & Nelson, 2003; Criss & Malmberg, 2008; DeLosh & McDaniel, 1996). This differential attention at encoding time is thought to make the memory trace more easily distinguishable from other memory traces, that is, the memory traces become comparably unique (Diana & Reder, 2006; Criss & Malmberg, 2008; Malmberg & Nelson, 2003). In addition to evidence from memory studies which limit the attentional resources available at study time (Diana & Reder, 2006), the idea that LF words require more attentional resources at encoding is supported by the observation that lexical access is slower for LF words compared to HF words (this idea was discussed in the previous section). The possibility that attention is elevated to LF words at study time is not mutually exclusive with the possible sources for a LF word's uniqueness listed before. In the contrary, it is

well conceivable that LF words attract more attention *because* they have more unusual features or have been previously encountered in a limited number of contexts.

Note that most of the theory and empirical evidence reviewed here, suggesting that LF words are more unique, stems from studies on recognition memory. However, if it is a characteristic of the stimulus material for LF words to be unique, the same argument can be made for the word frequency effect in recall and its abolition in mixed study lists. Although most theories attempting to explain word frequency effects on recall are centered around differential facilitation of elaborative processes as discussed in the following subsection, some authors do suggest that enhanced perceived uniqueness of LF words in mixed lists, and therefore enhanced item-based processing of LF words, contributes to the reversal of the word frequency effect on recall (e.g. DeLosh & McDaniel, 1996).

#### **1.2.4 Do HF Words Enhance Item-to-Item Elaborative Processes?**

One commonly cited observation is that HF words tend to have more meanings (Glanzer & Bowles, 1976), and lead to more associations than LF words (Cofer & Shevitz, 1952). A related finding suggests that HF words have fewer pre-existing associative links *from* (but not *to*) other words in lexical memory (Nelson & McEvoy, 2000). Thus, pure lists of HF words are often believed to have more pre-existing associations with each other than pure lists of LF words, and mixed lists score in between. Activating these associative links at study time can then enhance later retrieval success. Therefore, in pure lists, HF words should facilitate elaborative processes such as semantic elaboration or formation of item-to-item associations at study time and thereby lead to enhanced recall compared to LF lists (Mandler et al., 1982). In mixed lists of LF and HF words, the level of item-to-item associations is intermediate, leading to enhanced recall for LF words, but attenuated recall for HF words, compared to pure lists (Tse & Altarriba, 2007; Hulme et al., 2003). It is readily apparent that this account implicates elaborative processes between items in a list, rather than individual item-based processes, as the basis for word frequency effects on recall.

Another theory that implicates elaborative, or organizational, processes as the basis for word frequency effects on recall, but that differs with respect to the specific processes, is the *order encoding hypothesis*. This hypothesis states that LF words require more attentional resources to be encoded episodically (see above), but that this enhanced item encoding comes at a cost of reduced encoding of presentation order (DeLosh & McDaniel, 1996; Merritt, DeLosh, & McDaniel, 2006). HF words,

in turn, do not require as much attention directed toward the item-specific information, and therefore leave more resources available to enhance encoding of presentation order, a process that may be attributed to elaborated processes between list items or between the study item and its temporal context.

When considering the suggestions of (1) enhanced item-specific processing of LF words in lists of HF words due to their greater uniqueness or (2) enhanced elaborative processes due to the presence of HF words as explanations for word frequency effects on recall, it is worth noting that several authors explicitly suggest that *both* of these processes should contribute to the effects on recall (DeLosh & McDaniel, 1996; Tse & Altarriba, 2007). Hence, it should be kept in mind that although both of these hypotheses are tested in the current study, they are not tested *against each other*, but independently of each other. In other words, it is possible that both hypotheses will be supported by this study.

As outlined above, we will test the two hypotheses by investigating which brain potentials are pronounced and correlated with recall success for lists of HF and LF words with different list compositions. The most relevant brain potentials for this study are introduced in the next section.

### **1.3 The P300, the Frontal Slow Wave, and the N400**

When investigating brain processes associated with memory encoding of discrete events, the technique of event-related potentials (ERPs) with its excellent temporal resolution can provide useful insights, as very similar behavioral patterns can be associated with quite different brain wave characteristics (Otten & Donchin, 2000, for example). The ERP is the record of the responses, over a period of several hundred milliseconds, of various brain structures to discrete events. Because the ERP is time locked to the eliciting event it can be extracted by means of signal averaging from the ongoing electroencephalographic activity (Luck, 2005).

The main goal of the current study is to investigate the relationship of an ERP component known as the *P300* to subsequent memory for physically distinctive words and for words of different frequencies in the English language. The interest in this component stems from a large number of previous studies indicating its close correspondence to distinctiveness processing for items that stand out from their immediate context. Therefore, the first part of this section will review some characteristics and a theory about the functional significance the *P300*.

Two other ERP components, the *N400* and the *frontal slow wave*, are of interest for the current study. The *N400* can be considered an index of a different type of distinctiveness, namely distinctiveness with respect to semantic expectations based on the linguistic context. The functional significance of the frontal slow wave is rather unclear. It is often suggested that it reflects working memory based operations. This component has been found to correlate with memory in several previous ERP subsequent memory studies, and will therefore be investigated in the current study as well.

### 1.3.1 The P300 as a Measure of Distinctiveness

Of particular interest for the current study is the *P300* - an ERP component that is elicited by events that “resolve uncertainty”, as first reported by Sutton and colleagues (Sutton, Braren, Zubin, & John, 1965). The *P300* has a positive polarity and peaks at a latency of at least 300 ms, with a maximum at parietal recording sites. The *P300* is most commonly studied in the so-called *oddball paradigm* (Donchin, 1981). In this paradigm, two types of stimuli are presented in a random sequence. One type occurs frequently, and the other type occurs rather infrequently. The subject is engaged in a task that requires the classification of the stimuli according to the two types. Under these conditions, the infrequent stimulus elicits the *P300*. Note that the oddball paradigm and the Von Restorff paradigm share important characteristics such that in both tasks, an infrequent “oddball” item (or isolate) occurs, that the participant is likely perceive as distinctive.

The amplitude of the *P300* is negatively correlated with the frequency of occurrence of the rare stimulus. That is, the more rarely the infrequent stimulus occurs, the larger is the amplitude of the *P300* (Donchin, 1981; K. C. Squires, Wickens, Squires, & Donchin, 1976; Duncan-Johnson & Donchin, 1977). Furthermore, the shape of the *P300* is not necessarily determined only by the global frequency of the rare event, but rather by the subject’s *perception* of the frequency of occurrence at the time the stimulus is encountered (K. C. Squires et al., 1976; Duncan-Johnson & Donchin, 1977). Therefore, it has been proposed that the *P300* can be used as a measure of subjective distinctiveness (Donchin & Fabiani, 1991).

The *context updating theory* is one influential hypothesis on the functional significance of the *P300* that was developed in the early 80’s (Donchin, 1981). This hypothesis assumes that humans maintain a mental model – or schema – of all goal-relevant information about the current situation

and their environment. Whenever an event occurs that is novel or inconsistent with the schema, the model needs to be revised. The novel information has to be integrated into the model while taking into account information stored in long-term memory. According to the theory, whenever this context updating process occurs, a P300 is elicited. These cognitive processes that are supposedly associated with the P300 can be described as *strategic processing* (Donchin, 1981). That is, rather than merely reacting to the perceived event, these processes affect the way the subject will react to future events. The development of the context updating theory led to specific predictions about the consequences of the P300. One of these predictions is that if a large P300 is elicited by a distinctive event – indicating that a strong updating process of the schema occurs – this should increase the probability of later recall success (Donchin, 1981).

To explain this prediction more specifically, it is known that items that are distinctive from their context will elicit a P300. The amplitude of the P300, however, varies between the encounters of different distinctive items, that is, some distinctive items will elicit a larger P300 than others. As described above, the precise amplitude on a given trial can be considered a measure of the strength of the context updating process needed to integrate this novel information, or a measure of *subjective distinctiveness*. It is further known that behaviorally, distinctive items are typically more likely to be recalled (the Von Restorff effect). The prediction arising from the context updating hypothesis is that the amplitude of the P300 elicited by specific distinctive events – as an indication of how much the schema has to be updated to integrate this new information – should be correlated to whether or not the event will be successfully remembered in the future. In other words, trial-by-trial variability of the P300 – as it is found in an oddball paradigm or the Von Restorff paradigm, should correlate to the probability of subsequent recall. Evidence for precisely this pattern of results is reviewed in the section following the review of the slow wave and the N400.

### **1.3.2 Slow Waves as Index for Working Memory Processes**

Slowly increasing or decreasing potentials in the ERPs have first been observed in the mid-70's. These earliest reports indicated that a positive slow wave potential followed the P300 in different instances of the oddball task (N. K. Squires, Squires, & Hillyard, 1975; Duncan-Johnson & Donchin, 1977). In these studies, the slow positive wave showed a parietal maximum and was present only for stimuli that also elicit a P300. Parietally distributed positive (Bosch, Mecklinger, & Friederici,

2001) and negative (Ruchkin, Johnson, Canoune, & Ritter, 1991; Deldin, Deveney, Kim, Casas, & Best, 2001) slow waves have since been observed in various further studies. In addition to these parietally distributed slow waves, frontal negative (Bosch et al., 2001) and positivity (Ruchkin et al., 1991) slow waves, sometimes with a left- or right lateralized maximum, have been discovered in different experimental paradigms. To this date, the precise functional significance of the slow waves with different spatial distributions remains controversial. A variety of cognitive processes have been proposed, such as elaborative processes performed on stimuli (Deldin et al., 2001), resource allocation and memory-related processes (Ruchkin et al., 1991) or working memory related processes (Deldin et al., 2001; Bosch et al., 2001). It is not clear, however, whether slow waves with different spatial distributions elicited in different experimental paradigms serve similar cognitive functions, or whether they are entirely separable from each other. Therefore, the following discussion will focus only on slow waves that are maximal in frontal electrodes, because this is the spatial distribution relevant for the current study.

Thus, due to the findings of previous subsequent memory studies utilizing a similar paradigm as the current study (see next chapter), a frontally distributed slow wave is of biggest interest. Based on its frontal spatial distribution, it is likely that this frontal slow wave component originates in prefrontal areas. The prefrontal cortex has been repeatedly implicated in cognitive control processes, among others in working memory processes (for a review, see Miller, 2000).

Ruchkin and colleagues reported prefrontally distributed slow waves in mental arithmetic and mental rotation tasks, which increased in amplitude when the task difficulty increased (Ruchkin et al., 1991). Bosch and colleagues found that a similar frontal slow wave varied dependent on whether verbal, object-based or spatial information was to be held in working memory (Bosch et al., 2001). Furthermore, Münte and colleagues conducted an ERP study in which sentences associated with high and low working memory demand were presented (Münte, Schiltz, & Kutas, 1998). A slowly increasing left anterior negativity was larger when working memory demand was high compared to when it was low. This effect was modulated by individual subject's working memory capacity.

These and other studies are consistent with the suggestion frontally distributed slow waves are associated with some aspect of working memory operations. It is worth noting, however, that frontal slow waves have also been suggested to play a role in other cognitive operations, among others in retrieval of information from long term memory (Mecklinger, in press) or emotion processing

(Diedrich, Naumann, Maier, & Becker, 1997). Therefore, the precise cognitive processes manifested in the frontal slow wave remain to be elucidated.

### **1.3.3 The N400 as an Index of Semantic Integration**

An ERP component that has been linked to the effort necessary to semantically integrate newly encountered words into the context is the *N400* (Kutas & Hillyard, 1980). Interestingly for this study, this component can also be considered as a measure of semantic expectancy violation and may therefore be considered as reflecting a different type of “distinctiveness” (Fabiani, 2006). The *N400* has a negative polarity, and usually peaks around 400 ms after the semantically unexpected stimulus. In the original study, the *N400* was elicited by the final word in a sentence, and was large for sentence endings that did not meet semantic expectations (for example, “He spread the warm bread with socks.”) (Kutas & Hillyard, 1980). Subsequent studies indicated that the *N400* is smaller for words that are in the same semantic category as the expected word, while it is larger for words that are of different semantic categories (Federmeier & Kutas, 1999). This finding is consistent with the notion that the *N400* is a measure of the effort needed to integrate a word into its semantic context. While the original studies on the *N400* used entire sentences, later studies found that an *N400* can be elicited in the absence of a sentential context. That is, an *N400* is also elicited by the second word in a pair when this word is semantically unrelated to the first one (cf. (Kutas, Van Petten, & Kluender, 2007)).

It is interesting and important for the current study that a few studies indicated that word frequency is negatively correlated with the magnitude of the *N400*. Thus, LF words have been reported by several authors to elicit larger *N400* amplitudes than HF words (Van Petten & Kutas, 1990; Barber, Vergara, & Carreiras, 2004; Allen, Badecker, & Osterhout, 2003; Rugg, 1990). This effect appears to be especially powerful in the absence of strong semantic expectations, developed for example from the sentential context, which may constrain semantic expectations to specific subsets of words.

## **1.4 Correlations of ERP Amplitudes and Memory for Distinctive Stimuli**

A sequence of ERP studies utilizing the Von Restorff paradigm investigated the prediction that *P300* amplitudes are correlated with subsequent memory success. In the first study (Karis, Fabiani,

& Donchin, 1984), participants studied lists of 15 words and immediately after each list completed a free recall test. In most lists, one word was printed in a larger font size than all other words (“size isolate”). Between subjects, different behavioral patterns were found, such that one group showed a Von Restorff effect but a rather poor overall memory performance, while another group did not show a Von Restorff effect but overall performed much better, and a third group scored in between. The authors found that the groups differed with respect to the memorization strategies they had used at study time. The first group had silently repeated each word several times after its presentation (“rote memorization”). The high-performing group with no Von Restorff effect, in turn, had used elaborative strategies such as connecting the study words into sentences. For all groups, a clear P300 was elicited by the isolates, but only for the rote memorization group was its amplitude correlated with subsequent recall. For the group that had used elaborative strategies, another ERP component – the *frontal positive slow wave* – was pronounced, and this component was larger for subsequently recalled than subsequently not recalled words. Karis and colleagues concluded that the P300 as an index of early cognitive processing of a stimulus is only related to subsequent recall when no additional, elaborative processing of the word occurs afterwards.

Subsequent studies further investigated whether the correlation between P300 amplitude and subsequent recall is tied to the absence of elaborative memorization strategies. In one study (Fabiani, Karis, & Donchin, 1986), subjects performed an oddball tasks in which they classified names according to their gender. At study time, participants were not informed that a free recall test would follow, so that participants were unlikely to use elaborative strategies. Infrequent items associated with a large P300 at study time were more likely to be recalled than items associated with a smaller P300. This finding corroborates the original idea that when elaborative memorization strategies are not used, P300 amplitude is correlated to subsequent recall. Note that in this study, there was no evidence for a frontal slow wave, further supporting the notion that this ERP component may be associated with elaborative processes (Fabiani et al., 1986).

Another study used a similar paradigm as Karis and colleagues on 11-year old children (Fabiani, Gratton, Chiarenza, & Donchin, 1990). The authors found that children showed strong Von Restorff effects and a strong correlation between the P300 and subsequent recall. As children are not likely to engage in elaborative processing strategies, the authors reasoned, this finding again indicates that in the absence of elaborative strategies, the P300 is correlated with subsequent recall.



The goal of a fourth study in this series (Fabiani, Karis, & Donchin, 1990) was to directly investigate, within subjects, the interactions between P300, subsequent memory and rehearsal strategies. For each list, participants were instructed to use one of the two types of strategies. When participants used rote memorization strategies, there was a strong Von Restorff effect as well as a correlation between P300 amplitude elicited by distinctive items and subsequent recall. When participants used elaborative strategies, a frontal positive slow wave was prominent and correlated to recall. The results confirm that when rote memorization strategies are used, the P300 is related to subsequent recall, but when elaborative strategies are used, a relationship to recall is found for the frontal slow wave.

The next study investigated the impact of semantic and physical distinctiveness on the relationship between P300 and recall (Fabiani & Donchin, 1995). Subjects were instructed to use rote memorization strategies to study word lists, while at the same time either completing a physical task (judging the font size of each word), or a semantic task (judging whether or not the item is an actual word or a non-word). In each list one item was distinct with respect to semantic class (“semantic isolate”), and one word was distinct with respect to its font size (“size isolate”). For both groups, the P300 in response to physical isolates was correlated to subsequent success of recall. The semantic isolate elicited a significant P300 only in the group completing the semantic task, and this P300 was correlated to subsequent recall as well. Furthermore, the authors also found that semantic isolates elicited an N400. However, its amplitude was not correlated with subsequent recall (Fabiani & Donchin, 1995).

The latter study also used a recognition test, which indicated that the relationship between P300 and memory can be found in a recognition paradigm if the memory test uses the original presentation style during test and thereby reinstates the study context of study. Fabiani and Donchin additionally investigated how likely it is that a distinctive study item is written down first or last in the recall phase, as opposed somewhere to within the bulk of non-distinctive study items. They found that semantic isolates were most likely to be recalled first, whereas physical isolates are most likely to be recalled last. This finding indicates that these two types of distinctiveness are processed differently, while both types of distinctiveness seem to be available as aids for the organization of the list items in memory (Fabiani & Donchin, 1995).

Another study in this series investigated how the relationship between P300 and subsequent recall is affected by the manner in which physical distinctiveness is induced (Otten & Donchin, 2000). Thus, the correlation may be due to a general process for all distinctiveness attributes, or it may be only present when distinctiveness is integral to the study item. Integral distinctiveness was manipulated by font size (“size isolate”), while non-integral distinctiveness was manipulated by drawing a near, or a far frame around the word. The results indicated that behaviorally, integral distinctiveness had the same enhancing effect on recall performance as non-integral distinctiveness. However, the ERP data indicated that the size isolate elicited large P300 components, and only for this type of isolate was the amplitude of the P300 correlated with subsequent recall. The words surrounded by a far frame, in turn, showed a smaller P300 that was not correlated to subsequent recall. Instead, the frontal slow wave showed a correlation to memory. In an attempt to explain this finding, the authors speculated that the P300 may be more pronounced and the amount of context updating may be more strongly related to subsequent recall, when the distinctiveness feature has to be necessarily processed during encoding (that is, it has to be integral to the study item).

One shortcoming of all studies reviewed so far is that they all used only 10 or less electrodes due to the limited technology at the time they were conducted. Two subsequent studies (Lian, Goldstein, Donchin, & He, 2002; Ding, Goldstein, Lian, Donchin, & He, 2002) used recordings from a 129-sensor dense electrode array in a paradigm analogous to Karis et al. (1984). They applied a cortical imaging technique to map the ERP data onto likely cortical regions of origin. The results indicated that for the rote memorization condition, the ERP subsequent memory effect was mapped onto the left inferior prefrontal cortex, the left medial temporal lobe and the left parietal lobe. For the elaborative strategy condition, activations in the right parietal and left prefrontal lobes differed in magnitude between subsequently recalled and subsequently not recalled items.

The combined results of these studies led to the suggestion that processing of distinctive items in a Von Restorff paradigm can be described in three phases. First, during an initial, feature-driven encoding of the distinctive item, a P300 is elicited whose amplitude varies according to the amount of context updating that is elicited. This context updating process involves marking the distinctive attribute towards the memory representation, such that the stronger the updating process (and therefore the larger the P300), the stronger the distinctiveness attribute is marked. This feature-driven representation is created independently of memorization strategy. During the second phase, the use

of different rehearsal strategies operates on the memory representation. If participants use memorization strategies like forming sentences with the words, an elaborated representation is created containing the results of this process. This elaborated representation involves forming associations between study items. The theory leaves open whether the elaborated representation is independent of the feature-driven representation or whether the original representation is altered. The third phase is the retrieval phase, in which subjects are able to use different kinds of retrieval aids. For rote memorizers, the distinctive attribute of the isolate can be effectively used as a cue for retrieval. If the original representation was elaborated on in the second phase, however, the distinctiveness attribute may be less beneficial as a retrieval cue, because the organization induced by the elaboration is stronger and more efficient for retrieval. It is for this reason that elaborative rehearsal at time of study is associated with a reduction, or absence, of the correlation between the P300 and subsequent recall (Fabiani & Donchin, 1995; Donchin & Fabiani, 1991).

Related to this theory, one possibility is that the updating process associated with the P300 is correlated to recall when the actual study episode is retrieved. The frontal slow wave, in turn may be related to subsequent memory when working memory operations in the frontal lobe are activated, which enable relational processing providing effective cues for retrieval (Donchin & Fabiani, 1991). Note that other authors suggest the opposite pattern: that the frontal slow wave is correlated with memory when contextual detail is retrieved while the P300 shows a correlation when no contextual detail is retrieved (Kim, Vallesi, Picton, & Tulving, 2009). Therefore, the nature of the dissociation between correlations of the two brainwaves to memory remains to be elucidated.

Many studies in other laboratories provide converging evidence for more positive-going ERPs for subsequently remembered compared to not remembered events, either in the P300- or in the slow wave range. Thus, several other studies have found correlations between P300 amplitudes (Paller, Kutas, & Mayes, 1987; Azizian & Polich, 2007; Voss & Paller, 2009), slow wave amplitudes (Mangels, Picton, & Craik, 2001; Otten, Sveen, & Quayle, 2007), or both under different conditions (Kim et al., 2009; Wiswede, Rüsseler, & Münte, 2007) to subsequent memory in different experimental designs, such as traditional study of sequentially presented word lists, or paired associates tasks. While these studies did not specifically investigate the effects of physical isolation on memory, they do provide converging evidence supporting the generalizability of the claims about cognitive processes associated with P300 and slow wave and their relationship to subsequent recall.

As discussed above, the term “distinctiveness” is commonly used to explain different phenomena. The current study will build on previous studies on the Von Restorff paradigm and investigate how physical isolation compares to isolation due to deviant word frequency. That is, we will investigate whether words of deviant frequency elicit a P300 and a frontal slow wave, and furthermore which one of the elicited components is correlated with subsequent recall. Thereby, we will determine whether the effects of the two types of distinctiveness on memory are supported by similar, or separable, neuro-cognitive processes.

Therefore, of interest for the current study are previous studies on the effects of word frequency on recall memory, as well as the corresponding brain wave correlates. A review of the ERP correlates of word frequency effects on memory follows next.

### **1.5 ERP- and Other Neuroimaging Studies on Word Frequency Effects**

This section will provide a review of neuroimaging studies on the effects of word frequency on neuro-cognitive processes, as well as their correlations to subsequent recall. First, it is relevant whether LF words can under certain circumstances elicit a P300, which should be the case if LF words are perceived as distinctive in a similar way as items that physically stand out.

Polich and Donchin designed an oddball paradigm in which the participants judged each stimulus in a sequence by whether it is a word or a pronounceable non-word (*lexical decision task*) (Polich & Donchin, 1988). The probability of words and non-words occurring on any given trial was manipulated between experimental blocks. Of the words, 50% were LF words and 50% were HF words in every block. The findings indicated that both LF and HF words elicited a P300, which was negatively correlated with the probability of an actual word occurring on any given trial. Furthermore, the P300 elicited by LF words was smaller in amplitude and showed a longer latency compared to the P300 elicited by HF words. The authors interpreted this finding in terms of the cognitive capacity needed for processing common and uncommon words: uncommon words require more processing capacity and -time for their identification leading to the reduction in P300 amplitude and the increase in P300 latency (Kutas, McCarthy, & Donchin, 1977). Note that although word frequency had an impact on characteristics of the P300, the probability of an actual word to occur within a block was a much stronger predictor of P300 amplitude (Polich & Donchin, 1988).

This finding suggests that a stronger P300 is not necessarily elicited by LF words when compared to HF words. Thus, considering LF words as distinctive without regard to their context does not appear valid when subjective distinctiveness is defined as the processes manifested in the P300 (although other definitions of distinctiveness may still hold for LF words). The more important question of this study, however, is whether LF words are identified as distinctive when they occur with a lower probability in sequences of HF words. Thus, it is possible that the direct contrast with HF words “brings out” the LF words’ distinctiveness. A study systematically investigating this issue using the oddball paradigm or a similar design is not known to us. However, several previous studies have addressed related questions and therefore are worth reviewing.

A series of studies by Rugg and colleagues investigated ERPs elicited by LF and HF words in lexical decision paradigms in which words were repeatedly presented (Rugg, 1990), or in which a recognition test was administered in a second phase (Rugg, Cox, Doyle, & Wells, 1994; Rugg & Doyle, 1992). The first study suggested that repeated presentations of LF words elicited larger parietal positivities in the P300 latency range compared to first presentations, a finding that was not obtained for HF words (Rugg, 1990). Similarly, the subsequent studies suggested that the “old/new effect” (i.e. more positive going ERPs for previously studied words than for new words) is large for LF words but absent for HF words (Rugg et al., 1994; Rugg & Doyle, 1992). There is a methodological difficulty when interpreting these findings such that all three studies measured ERP amplitudes by averaging over large intervals of durations between 100 ms and 400 ms. With this type of analysis, one cannot clearly disentangle different ERP components, preventing any definite conclusions about the componential nature underlying their findings. These studies do, however, provide some indications that P300 amplitudes elicited by repeatedly presented words may be larger for LF words than HF words. A subsequent memory analysis was not the goal of any of these studies, preventing any conclusions about possible subsequent memory effects.

Experiments investigating subsequent memory effects for LF and HF words are most interesting for the current study. The most important question is whether amplitudes of the P300 or the frontal slow wave are correlated with subsequent recall and under which conditions. However, so far, only two studies have directly investigated ERP subsequent memory effects elicited by LF and HF words in a recall memory paradigm (Fernandez et al., 1998; Guo, Zhu, Ding, Fan, & Paller, 2004).

The earlier study's objective was to manipulate distinctiveness by word frequency, such that LF words were a-priori considered distinctive (Fernandez et al., 1998). Under rote-memorization instructions, participants studied lists of 15 words presented sequentially, while ERPs were recorded. The word lists either contained either only LF or only HF words (blocked lists), or about half of the words in a list were LF and half of the words were HF words (mixed lists). The ERPs were collapsed across list types, because according to the authors the two list types did not elicit differences in ERPs. Interestingly, a "late positive complex" (that is, a P300) was elicited by LF words only, which was used as an indication for the success of the distinctiveness manipulation. The amplitude of the P300 was, however, not correlated with subsequent recall. The explanation for the discrepancy between the findings of this study and the series of studies on the P300 and Von Restorff paradigm experiments is unclear. It is conceivable that the P300 in this study was driven by only a small number of LF words that the subjects did not know and therefore stood out from the list. However, these words then did not lead to a recall advantage due to the lack of semantic content (since unknown words may be perceived as non-words). Fernandez and colleagues did find two subsequent memory effects, one that was present for both HF and LF words, and one that was only present for HF words. Due to the use of subtraction techniques and a lack of further investigating the componential nature of their findings, it is difficult to interpret which components were involved in these subsequent memory effects.

Another study investigating subsequent memory effects for words of different frequencies manipulated encoding task in a study-test recognition memory paradigm (Guo et al., 2004). In this study, subsequently remembered LF words were associated with a larger centrally distributed positivity in the P300 time window compared to not recognized LF words. Such a subsequent memory effect was not present for HF words. Again, one complication of this study is the use mean amplitude measures over time windows over 200 ms, preventing definite conclusions about the componential nature of these effects. However, this study's suggestion - the presence of a correlation of P300 to subsequent memory for LF words - appears to contradict the results obtained by Fernandez and colleagues. There are a few differences in methodology between the two studies possibly contributing to the different results, among others the use of a recognition test by Guo and colleagues as opposed to a recall test by Fernandez and colleagues. Overall, the conflicting results between the two studies imply that there is a need for further studies clarifying under which conditions P300 and frontal

slow wave amplitudes are correlated with subsequent recall for word frequency effects. Another question left open by both of these studies is the impact of the list composition on word frequency and subsequent memory effects of ERPs. The latter is one of the goals of the current study.

As discussed above, another ERP component that varies with word frequency is the N400, whose amplitude is inversely correlated with word frequency (Van Petten & Kutas, 1990; Barber et al., 2004; Allen et al., 2003; Rugg, 1990). To our knowledge, no studies have investigated a potential correlation between its amplitude and subsequent memory in a paradigm comparing LF and HF words. However, Fabiani and Donchin did not find a correlation between N400 and subsequent recall for words that belonged to a deviant semantic category (Fabiani & Donchin, 1995). Therefore, in our study we did not expect any correlations between N400 and recall when manipulating word frequency.

Another method of studying the brain's response to stimuli, which is also commonly used in subsequent memory paradigms, is functional magnetic resonance imaging (fMRI). The temporal resolution of this method is rather poor compared to ERPs, although it bears a great spatial resolution. Another property worth noting is that fMRI measures blood oxygenation levels in different parts of the brain, which is an indirect measure of brain activity. ERPs, in turn, measure electrical potentials emanated from neuron assemblies in the cortex. It is important to keep in mind these differences when integrating findings from ERP and fMRI studies. Nevertheless, useful insights can be gained if findings from fMRI and ERP studies converge on the same, or analogous, conclusions.

One previous study investigated subsequent memory effects of LF and HF words with event-related fMRI (Chee, Westphal, Goh, Graham, & Song, 2003). The results of this study indicated that LF words in an incidental learning paradigm were associated stronger blood oxygenation in left prefrontal regions than HF words. Furthermore, this prefrontal activation was stronger for LF words that were correctly recognized in a subsequent memory test compared to subsequently not recognized LF words. Keeping in mind the precautions discussed above for comparing ERP and fMRI data, one can speculate that the subsequent memory effect in prefrontal areas reported by Chee and colleagues corresponds to the slow wave activity observed in previous ERP studies. It is important to note, however, that the study by Chee and colleagues used a very different experimental design, such that recognition memory was tested, and the memory test occurred unexpectedly.

Overall, previous results indicate that LF words do not necessarily elicit a P300 component, but can sometimes do so depending on the details of the experimental design, such as the LF words' immediate context. Few studies have investigated subsequent memory effects of LF and HF words, and those that did provided conflicting results. It is therefore unclear up to this point, whether the P300 or the frontal slow wave are correlated with subsequent recall when word frequency is manipulated, and how the precise experimental conditions affect these correlations.

## **1.6 Objectives and Hypotheses for the Current Study**

The main goal of this study is to determine which ERP components are correlated with subsequent recall for LF and HF words, when they occur in lists of mostly same-frequency words and when they occur isolated in lists of opposite-frequency words. The obtained correlation between ERP components and subsequent recall will be compared to the subsequent memory effects for physically distinctive words to clarify whether effects of “primary distinctiveness” and “secondary distinctiveness” on recall refer to similar neuro-cognitive processes under certain conditions. Specifically, the question is whether the reversal of word frequency effects on recall when mixed lists are studied is due to enhanced perceived distinctiveness of LF words in lists of HF words (comparable to processes elicited by physical deviance) or whether it is due to enhanced between-item elaborative processes.

The study focuses on two ERP components: the P300 and the frontal slow wave. Previous studies indicate that the amplitudes of these two components are associated with subsequent recall under different conditions – the P300 when items are distinctive and when participants do not use elaborative memorization strategies; and the frontal slow wave when elaborative strategies are used and for non-distinctive items. The theoretical implication of these previous studies is that if the outcome of a memory updating process elicited by subjectively distinctive items is effectively utilized for retrieval, the P300 is to correlate to subsequent recall. If, in turn, the outcomes of elaboration between list items is most effectively used at retrieval, the frontal slow wave is correlated with recall.

Therefore, if our results show that LF words isolated in lists of HF words elicit a P300, and if its amplitude is correlated to subsequent recall, we conclude that the effects of physical distinctiveness and distinctiveness due to a low word frequency on memory share underlying cognitive processes under the given conditions. That is, a memory updating process elicited by physically deviant



words and by unexpected LF words is what causes the enhancement in recall rates. The reversal of word frequency effects on recall, in this case, could in part be explained by enhanced subjective distinctiveness of LF words when occurring in lists of HF words.

If, in turn, the frontal slow wave is correlated to subsequent recall for isolated LF words, we will conclude that outcomes of enhanced elaborative processes at study due to the presence of HF words in the same list are used as retrieval aids at the time of the test. In this case, the conclusion would be that subjective distinctiveness does not play a major role in the reversal of word frequency effect in mixed lists, but that facilitation of inter-item elaboration due to the presence of HF words is the cause.

An important control condition in the current study is composed of the HF words isolated in lists of LF words. As outlined above, LF words can be considered distinctive with respect to previous experiences, while HF words are not. If our findings are indeed due to this type of distinctiveness defined based on prior experiences (“secondary distinctiveness”), we should obtain effects for LF words isolated in lists of HF words, but not for HF words isolated in lists of LF words. If, however, comparable effects are obtained for isolated LF words *and* isolated HF words, the findings may in fact be another instance of “primary distinctiveness” that is independent of how often a word has been experienced before but only dependent on whether the word frequency *differs* from the word frequencies of all words in the surrounding context.

In addition to the P300 and the frontal slow wave, this study also investigates the N400. While previous research indicated that in the absence of a constraining sentential context, the N400 is larger for LF than for HF words, to our knowledge there are no studies that resulted in correlations between N400 amplitude and subsequent recall. Therefore, we do not make any directed predictions of whether or not we will obtain an N400 for any of the word types and if we do, whether its amplitude will be correlated to recall.

A further goal of the current study is to replicate previous findings with a dense electrode array and a detailed investigation of the spatio-temporal characteristics ERP components with a spatio-temporal principal component analysis (PCA) (Dien, Spencer, & Donchin, 2003). Thus, one shortcoming of many of the studies reviewed above is the limited number of recording electrodes used due to limited technology available at that time. Otten and Donchin’s (2000) study used 10 electrodes, while the even earlier studies only used only up to 3 electrodes. These low numbers of

electrodes do not allow for specific conclusions about spatial distributions of ERP components. In the current study we record ERPs with the 128 EGI electrode net now available in the USF Psychophysiology lab, which in combination with the spatio-temporal PCA can provide more detailed information of the spatio-temporal characteristics of components of interest.

## **Chapter 2**

### **Methods**

#### **2.1 Participants**

Nineteen undergraduate students at the University of South Florida completed two experimental sessions each, for a total duration of approximately 5 hours. Data from five participants were excluded due to excessive artifacts in the ERP data, and data from one participant were excluded due to severe incompletion with the instructions. The remaining 13 participants were between 18 and 45 years old (mean age: 25.15), and all reported English to be their first language. Eight participants were female and all but one participant were right-handed. Participants received partial credit for an introductory psychology course in return for their participation.

##### **2.1.1 Incentives for Good Performance**

To encourage good performance, participants were informed before the first session that their performance in the memory portion of the experiment would be scored. They were told that they would receive one point for each word they correctly recalled. If on a given list they wrote down more than three words that had not been studied on the previous list, they would lose points. Each participant was given the option to receive an email with their own score as well as the average score of all participants once their performance was evaluated.

Furthermore, the participant who reached the highest score was determined after data collection was finished, and the highest score was rewarded with a bonus of \$100. It is worth noting that no participant was informed about names or any other personal information of the other participants who had taken part in the study.

## 2.2 Materials

### 2.2.1 Stimuli

The words in the memory experiment were classified into three groups: high-, low-, or medium frequency words. For this grouping, we combined the word frequency norms by Francis and Kucera (Francis & Kucera, 1982), with subjective familiarity norms (Toglia & Battig, 1987), because pilot studies indicated that both normed frequency and normed familiarity affected the results. While most authors who study word frequency effects use only one norm in their operationalization, word frequency and word familiarity norms can be thought of as measuring the same underlying phenomenon (Connine et al., 1990), so the combination of the two norms seemed appropriate.

We created three groups of stimuli: (1) Words of high normed frequency (more than 50 lemma occurrences in a million) as well as high normed familiarity (mean rating of at least 6 on a scale of 1 to 7); (2) words of low frequency (less than 10 occurrences per million) and low familiarity (mean rating between 3.5 and 5.5); and (3) words with middle scores of normed frequency (between 10 and 50 occurrences per million) and normed familiarity (mean rating between 5.5 and 6.1). For the purpose of simplicity, the first word type will in the following be referred to as high frequency (HF) words, the second word type will be referred to as low frequency (LF) words, and the third word type as medium frequency (MF) words. We included only open-class words (nouns, verbs and adjectives) that were between 3 and 8 letters long.

Note that there is a slight overlap in the normed familiarity of MF words and HF words, due to a scarcity of stimuli that qualified for each group. The critical word frequency manipulation, however, is between HF and LF words, while the MF words were used to create Von Restorff lists in order to replicate previous studies. Therefore, we did not judge this slight overlap in normed familiarity as problematic.

All stimuli were presented one at a time on a computer screen in a white font on a black background for a duration of 250 ms, with a gap of 2000 ms between stimuli, in which a fixation cross was shown. Words were displayed in 16 pt font for regular sized words and 22 pt font for larger words (size isolates), in Arial Unicode style.

### 2.2.2 Word Lists

We used three types of lists to test our experimental hypotheses, and in addition created a set of filler lists. Each list, regardless of list type, consisted of 15 words.

The design of the first list type, which we will call *Von Restorff lists*, was designed very similarly to previous studies (Otten & Donchin, 2000). Of the 15 MF words used in each Von Restorff list, one was displayed in a larger font size (“size isolates”) than all the other words (“size standards”). The size isolate occurred at a serial position between 6 and 10, determined randomly for each list.

The second and third list type, respectively, consisted primarily of either HF words or LF words (“HF standards” and “LF standards”, respectively). One word in each of these lists differed in frequency from the rest of the list. That is, each HF list contained one LF word (“LF isolate”), and each LF list contained one HF word (“HF isolate”). The isolates were randomly placed at a serial position between 6 and 10. Each participant encountered a total of 25 HF lists and 15 LF lists throughout the experiment. The unbalanced number of word lists between the list types occurred due to a scarcity of words fulfilling the LF word criteria.

The last list type was constructed as filler lists that were intended to keep participants from developing expectations for an isolate in each list. Thus, these types of lists consisted of a mixture of words from different frequencies and normed familiarities and did not contain any isolates. There were a total of 10 such filler lists throughout the experiment.

For each participant, the grouping of words into lists as well as the order of words within the lists was randomized. Thus, each word was drawn randomly from the pool of 300 MF words, 235 LF words, 365 HF words, or 150 filler words, respectively. Note that this randomization procedure also insured that a word that one participant viewed as an isolate would be encountered by another participant as a non-isolate. Each participant studied each word only once during the entire course of the experiment.

In addition to this randomization procedure of the individual words, the presentation order of the list types was determined randomly as well. The only restriction was that 14 HF lists, 8 LF lists, 12 Von Restorff lists, and 6 filler lists (a total of 40 word lists) were studied in the first session while 11 HF lists, 7 LF lists, 8 Von Restorff lists and 4 filler lists (a total of 30 word lists) were studied in the second session.

### **2.2.3 Recall Sheets**

To record the participant's recall memory, one recall sheet was available to each participant for each study list on a clipboard. Each recall sheet included 15 lines, numbered from 1 to 15, as well as a little bit of space for comments at the bottom of the page. The list number each sheet corresponded to was printed on top of the sheet. A pen holder with several pens and pencils was positioned on the table in the experiment room, with which the participant could fill out the sheets during each recall phase.

### **2.2.4 Individual Word Familiarity Questionnaire**

A word familiarity questionnaire was administered on a paper-and-pencil base. To insure that the ratings obtained from this questionnaire were comparable to Toglia and Battig's ratings, the instructions and the design of the questionnaire were constructed to be as similar as possible to theirs (Toglia & Battig, 1987). Because of the large number of stimuli used in this experiment and to control for sequential effects on the ratings, we constructed three versions of this questionnaire that each contained only a subsets of the words. Each of the three versions contained all LF words (since this was the word type for which subjective familiarity was most critical), about one third of the HF words and about a third of the MF words, resulting in 300 words listed in random order for each version. Each of the HF words and MF words was randomly assigned to occur in exactly one of the three versions of the questionnaire. Next to each word, a likert scale from 1 ("very unfamiliar") to 7 ("very familiar") was displayed in which the participant could circle their familiarity rating for the respective word. Each participant completed only one version of the questionnaire.

### **2.2.5 Stimuli and Design of the Oddball Task**

In a standard oddball task participants viewed a random sequence of an "X" and an "O" presented on the screen. The probability of an "X" occurring on a given trial was 0.8, and the probability of an "O" occurring was 0.2. Due to these probabilities, for the "X" we will in the following use the term "standard" and for the "O" we will use the term "target". Overall, the sequence contained 200 stimuli (about 40 "O"'s and 160 "X"'s). The participant's task was to respond to each "X" by a left button press and to the "O" by a right button press. The stimuli were presented in black font on a white background in 100 pt Courier New font. Each stimulus was preceded by a fixation cross for

1000 ms and remained on the screen until the participant had given a response, or at most for 10 seconds.

### **2.3 Task and Procedure**

For each participant, each of the two experimental sessions lasted at most 2 1/2 hours and began with the usual preparations for the EEG recording. Then, participants were comfortably seated in the experiment room, at a distance of about 3 feet from the computer screen.

In the first session, after giving written informed consent and receiving instructions visually presented on the screen, the participant completed two practice lists. When any remaining questions about the task and procedure were answered, participants studied 40 word lists, each immediately followed by a recall phase. The start of the recall phase was indicated by the appearance of a large grey triangle displayed on the screen. Recall phases lasted for at least 45 seconds, but there was no upper time limit for completing this phase. The participant could finish the recall phase and start the next list with a button press whenever they were ready. The participant's task was to study each word using rote memorization strategies (that is, repeating each word silently after its presentation), and in the recall phase to write down every word they remember from the previous list in any order. A break was allowed after each set of 5 lists, in which the participant could relax for as long as they wished. During the breaks, the participant was reminded to use rote memorization strategies throughout the experiment.

The second session took place between three and seven days after the first session. First, participants received a shortened version of the instructions as a reminder. Then, they studied and recalled another set of 30 word lists under the same instructions as in the first session. After the last study - recall cycle was completed, they completed the oddball task. In this task, participants indicated the detection of an "X" by a left button press and the detection of a "O" by a right button press in each trial. After 4 practice trials, participants completed 200 experimental trials, which lasted up to 5 minutes. At the end of the second session, after the electrode net was removed, each participant was given unlimited time to fill out the subjective familiarity questionnaire. Most participants needed less than 10 minutes to complete the questionnaire. Finally, subjects were debriefed and questioned about the memorization strategies they had used, about whether they had any difficulties seeing the words on the screen, and whether they noticed that some words stood out from the lists.

## 2.4 EEG Recording and Analysis

The EEG was recorded with a 128 Electrical Geodesics system (EGI, Eugene, OR) in the Cognitive Psychophysiology lab at the University of South Florida. The EEG was digitized at a sampling rate of 250 Hz, and referenced to the central electrode (Cz). In an offline-analysis using Netstation software, the EEG was digitally low-pass filtered at a cutoff frequency of 20 Hz and bad channels were replaced by a mathematical procedure taking into account the signals of the channels adjacent to the bad electrode. Then, the continuous data were segmented from 400 ms before to 2000 ms after each stimulus. To correct for eye movement artifacts and eye blinks, we used an independent component analysis approach, as provided in Dien's ERP toolkit, version 1.3, based on MATLAB (Dien, 2010). Note that this eye movement correction procedure included a baseline correction using the average amplitude of 400 ms preceding the stimulus. Subsequently, each trial was visually inspected for remaining artifacts and the corresponding trials were excluded from any further analysis. In a final pre-processing step using Netstation software, the EEG was re-referenced to a linked mastoid reference.

All subsequent analysis was conducted using scripts and functions written in MATLAB, while utilizing several functions of the EEGLAB toolbox (Delorme & Makeig, 2004). To obtain the subject ERPs, we averaged over all artifact-free trials separately for each word type (HF, LF and MF), isolation type (isolate and non-isolate) and recall success. We only included trials from serial positions 6 to 10 to prevent confounds with primacy and recency effects. For illustration purposes, grand average ERPs were computed by combining all subject averages.

We conducted a spatio-temporal principal component analysis (PCA) (Dien et al., 2003) to identify and quantify ERP components, using version 1.23 of Dien's ERP toolkit (Dien, 2010). Promax rotations without the Kaiser normalization option were applied for both the spatial and the temporal analysis step. In the spatial step, a scree test indicated that 25 factors should be retained and rotated. For each spatial factor, the "virtual ERPs" were computed by averaging over the factor scores of all participants, separately for word types and subsequent recall success. In the temporal PCA step, which was conducted separately on the factor scores from each spatial factor, 8 factors were retained. The number of temporal factors to retain was again determined by a scree test and was held constant across spatial factors for the sake of simplicity. Thus, the input to the spatio-temporal PCA consisted of the ERPs from 129 electrodes and 600 time points (2400 ms), and from 13 participants



and 14 conditions (size standard, size isolate, HF standard, LF isolate, LF standard, HF isolate, and filler; each word type separately for subsequently recalled and subsequently not recalled items). Note that we included the filler words into the PCA in order to increase the number of observations. However, we did not analyze trials from this list type further since the filler lists did not control for normed word frequency or familiarity.

There were only 15 HF isolate trials for each participant, which included both recalled and not-recalled items. Due to this low trial count, plots of the HF isolate ERPs may appear rather noisy. More importantly, 3 participants had no artifact free trials for the HF isolate - recalled condition, resulting in three missing values for this condition. As a consequence, the grand average ERPs contain data from only 10 participants for recalled HF isolates, but data from 13 participants for the not recalled HF isolates. Furthermore, in order to be able to analyze as much of the information available for HF isolates, they were included in the PCA, and a matrix containing only zeros was submitted in place of any the missing ERPs. This procedure does not affect the results of the PCA, since a matrix containing only zeros does not induce any variance into the data.

## **2.5 Statistical Analysis**

Recall rates (the proportion of words correctly recalled in the immediately following recall phase) were computed for each participant, separately for each word type (HF, LF and MF words), for each isolation type (non-isolated items and isolates), as well as for each serial position to construct serial position curves. Von Restorff effects were analyzed by comparing the recall rates for size isolates to the recall rates of non-isolates within the same lists at serial positions 6 to 10, using two-sided paired samples t-tests. Word frequency effects at serial positions 6 to 10 were investigated with a 2 (word frequency) x 2 (isolate vs. standard) repeated measures ANOVA. For the behavioral analysis, all significance levels were chosen at 0.05.

All statistical analysis of ERP components was conducted on the factor scores obtained through the spatio-temporal PCA. In a first analysis step, HF isolates were excluded from the statistics due to the three missing values for the HF isolate recalled condition. Repeated measures ANOVAs with the within subject factors word type (5 levels: size isolate, size standard, HF standard, LF isolate, LF standard) and recall success (recalled, not recalled) were then conducted to test for main effects and interactions. Significant interactions were investigated by post-hoc planned comparisons. For

the overall ANOVA, we selected and further analyzed main effects and interactions with p-values smaller than 0.1, while in post-hoc tests the chosen significance level was 0.05. Furthermore, for a few specific factors and conditions of interest, we conducted planned comparisons for one given word type, comparing subsequently recalled to subsequently not recalled trials in a two-sided paired samples t-test.

Subsequent memory effects for HF isolates were analyzed separately from all other word types. Due to the missing values we conducted two-sided independent samples t-tests for unequal sample sizes, including 10 observations for HF isolates that were later recalled and 13 observations for HF isolates that were later not recalled.

## Chapter 3

### Results

#### 3.1 Debriefing

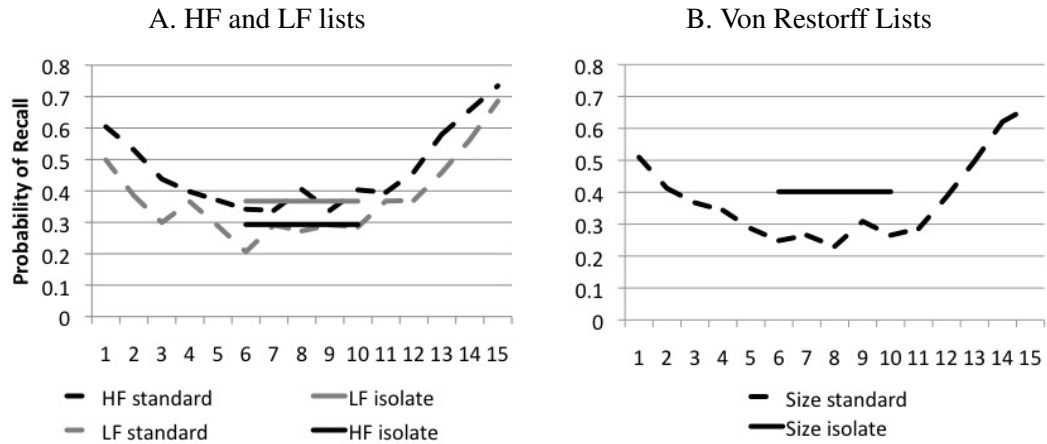
When asked whether they realized that some words stood out from the lists, every participant reported to have noticed that some words were printed in a larger font than the others, indicating that the manipulation of physical isolation due to font size was successful. Furthermore, most participants reported that they had tried to use rote memorization strategies. Two participants admitted to using strategies other than rote memorization occasionally during the experiment, and another two participants reported that they had used elaborative strategies most of the time (such as building sentences with the words or relating the words to each other in some other way).

#### 3.2 Behavioral Data

##### 3.2.1 Recall Rates

The serial position curve for the Von Restorff lists is presented in figure 1 B. The size isolates were significantly better recalled than the non-isolates at comparable serial positions (positions 6-10),  $t(12)=4.21$ ,  $p<.01$ . All but two participants showed increased recall rates for the size isolates than for the standards. Furthermore, by visual inspection of the serial position curves of the Von Restorff lists, it appears that for the standards, clear primacy and recency effects are present.

Figure 1 A shows the serial position curve for HF and LF lists. In the primacy serial positions (serial positions 1-5), HF standards were better recalled than LF standards,  $t(12)=3.52$ ,  $p<.01$ , which represents the traditional word frequency effect on recall when pure lists are studied. For the middle serial positions (6-10), the 2 (word frequency) by 2 (isolate vs. non-isolate) ANOVA revealed a significant interaction,  $F(1,12)=10.93$ ,  $p<.01$ . Post-hoc planned comparisons indicated that LF words were recalled better when they occurred isolated in lists of HF words compared to LF words



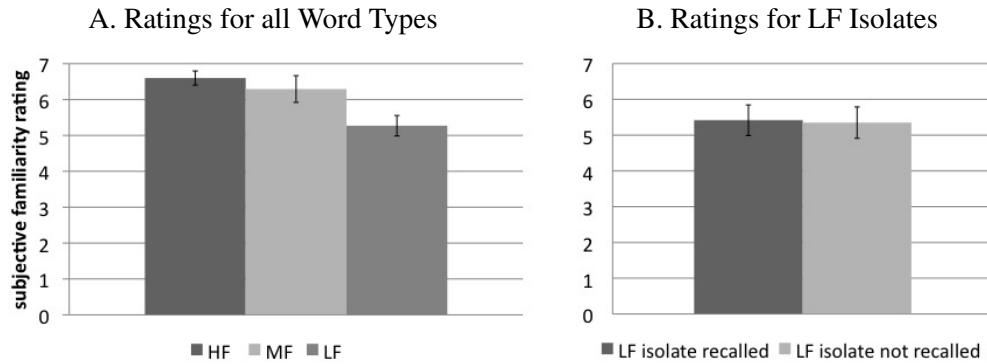
**Figure 1.:** Serial position curves. Probability of recall as a function of position in the list.

in LF lists. For HF words the reverse was true such that they were recalled *worse* when occurring in isolation in lists of LF words. Note also that within each list type the isolates were recalled at about the same level as the non-isolates. Thus, LF isolates were recalled about equally well as HF standards and HF isolates were recalled about equally well as LF standards. Technically, this finding differs from the Von Restorff effect, in which the isolates are recalled at a clearly *higher* level compared to the non-isolates within the same list.

In an additional analysis, we compared the recall rates for the non-isolated items in the Von Restorff lists (composed of MF words), LF lists and HF lists. A one-way ANOVA on the recall rates at serial positions 6-10 resulted in a significant main effect,  $F(2,24)=10.01, p<.01$ . Post hoc tests indicated that non-isolated HF words were more likely to be recalled than both non-isolated MF words and non-isolated LF standards, while MF and LF standards did not differ from each other.

### 3.2.2 Analysis of the Subjective Familiarity Questionnaire

To further investigate whether our manipulation of prior exposure using word frequency and familiarity norms was successful, we computed the average rating for each participant for words in the LF group, words in the MF group and words in the HF group. Note that each participant only rated a subset of the HF and MF words, but we assume here that those ratings are representative of the entire word type. The one way ANOVA on the subjective familiarity ratings revealed a significant main effect for word type,  $F(2,24)=21.79, p<.01$ . As can be seen in figure 2 A, HF



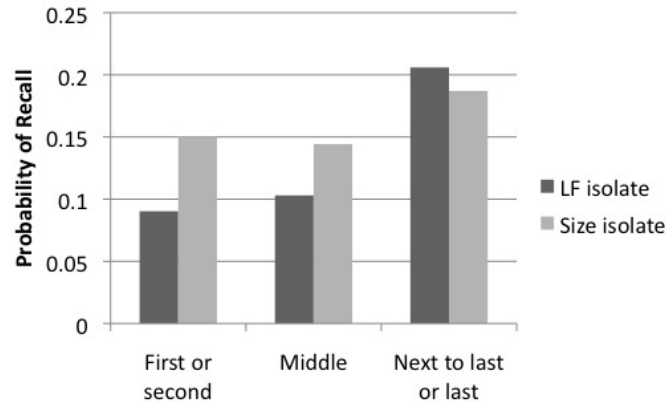
**Figure 2.:** Subjective familiarity ratings. A. Ratings for all word types, collapsed over isolates and non-isolates. B. Ratings only for LF isolates as a function of subsequent recall. Note that the actual words included in these categories are different for each participant. Error bars represent the standard error of the mean.

words were rated as more familiar than MF words, which in turn were rated as more familiar than LF words. This finding indicates that our manipulation of word type was successful such that our manipulation of prior exposure for our word types, as a group, matched the subjective experience of the participants. It should be noted, however, that several participants rated some LF words with a “1”, indicating that these words were “very unfamiliar or unknown” to them.

In addition, we analyzed the difference in familiarity ratings between LF words that were later recalled and LF words that were later not recalled. As can be seen in figure 1 B, there were no significant differences with respect to subsequent recall,  $t(12)=0.34$ , *ns*.

### 3.2.3 Analysis of Output Order

The graph in figure 3 shows the probability that a LF isolate or size isolate was written down at the beginning (output positions 1 or 2), the middle, or the end (next to last and last output positions) of the recall sheet, given that this isolate was successfully recalled. The probabilities do not add up to 1, since each probability is corrected for the number of words written down at the corresponding positions. Note that we do not report the corresponding analysis for HF isolates, because the number of recalled HF isolates was generally very low, causing high variability in the average probabilities for each output position.

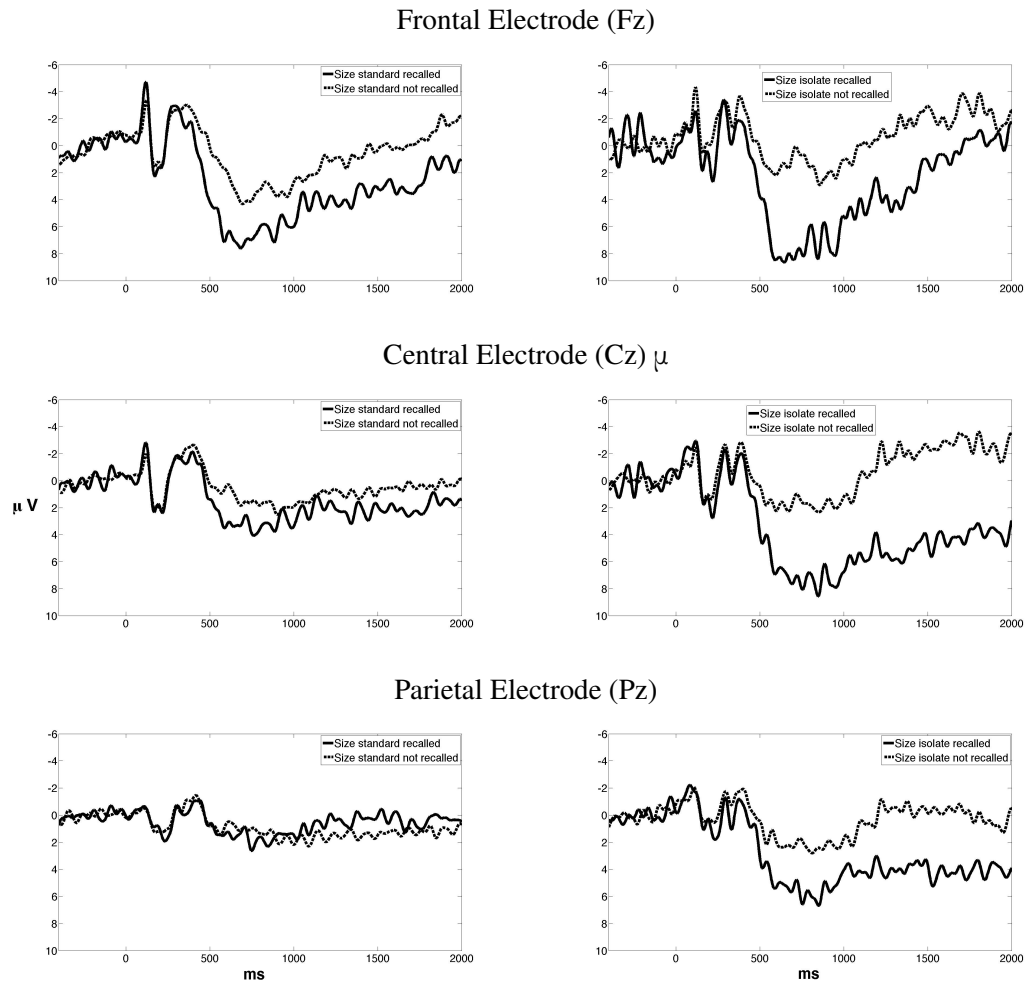


**Figure 3.:** Output order analysis. Probability of recall at specific output positions for size isolates and LF isolates, given that these were successfully recalled.

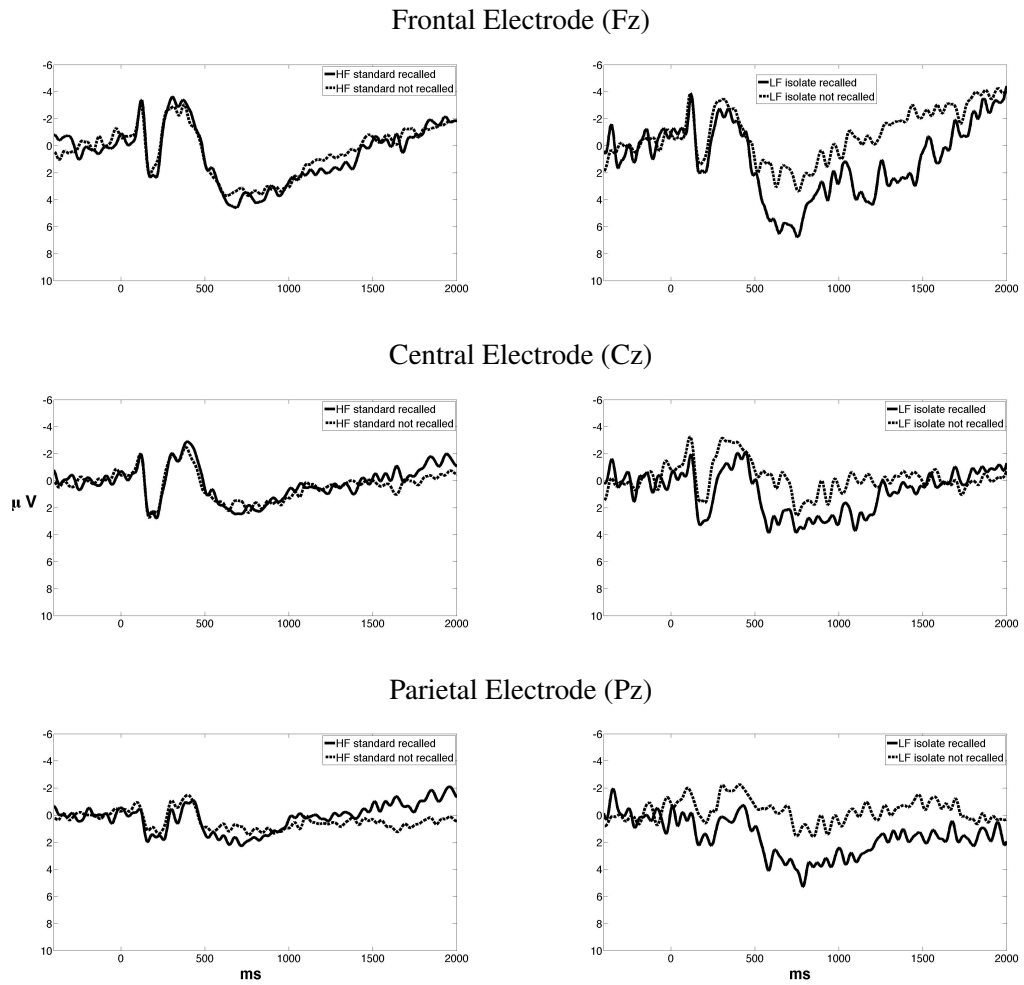
A one-way ANOVA on the output probabilities for LF isolates on serial positions 1 and 2, the middle output positions, and next to last or last output positions revealed a significant main effect,  $F(2,24)=3.52$ ,  $p=.05$ , indicating that LF isolates were more likely to be written down on one of the last two output positions, compared to middle or beginning serial positions. A similar analysis on the size isolates did not reveal a significant main effect for output position,  $F(2,24)=.55$ , *ns*.

### 3.3 Event-Related Potentials

Grand average ERPs are shown in figure 4 for the Von Restorff lists, in figure 5 for HF lists with LF isolates, and in figure 6 for LF lists with HF isolates. Visual inspection of the ERPs for the size isolates and the LF isolates indicates that a parietally distributed positivity, peaking around 700 ms, is larger for subsequently recalled compared to not recalled words (figures 4 and 5, right panels). The same is not apparent in the ERPs for HF isolates (figure 6, right panel) or any type of non-isolated word (figures 4, 5 and 6, left panels). The morphology and spatial distribution of this parietal positivity indicates that it represents an instance of the P300. In addition, a more frontally distributed sustained positivity, possibly an instance of the frontal slow wave, seems to distinguish between subsequently recalled and not recalled size isolates, LF isolates, and to a lesser degree, size standards (figures 4 and 5). Furthermore, for the HF isolates, a negative component maximal at 400 ms after the stimulus seems to distinguish between recalled and not recalled words (figure 6,

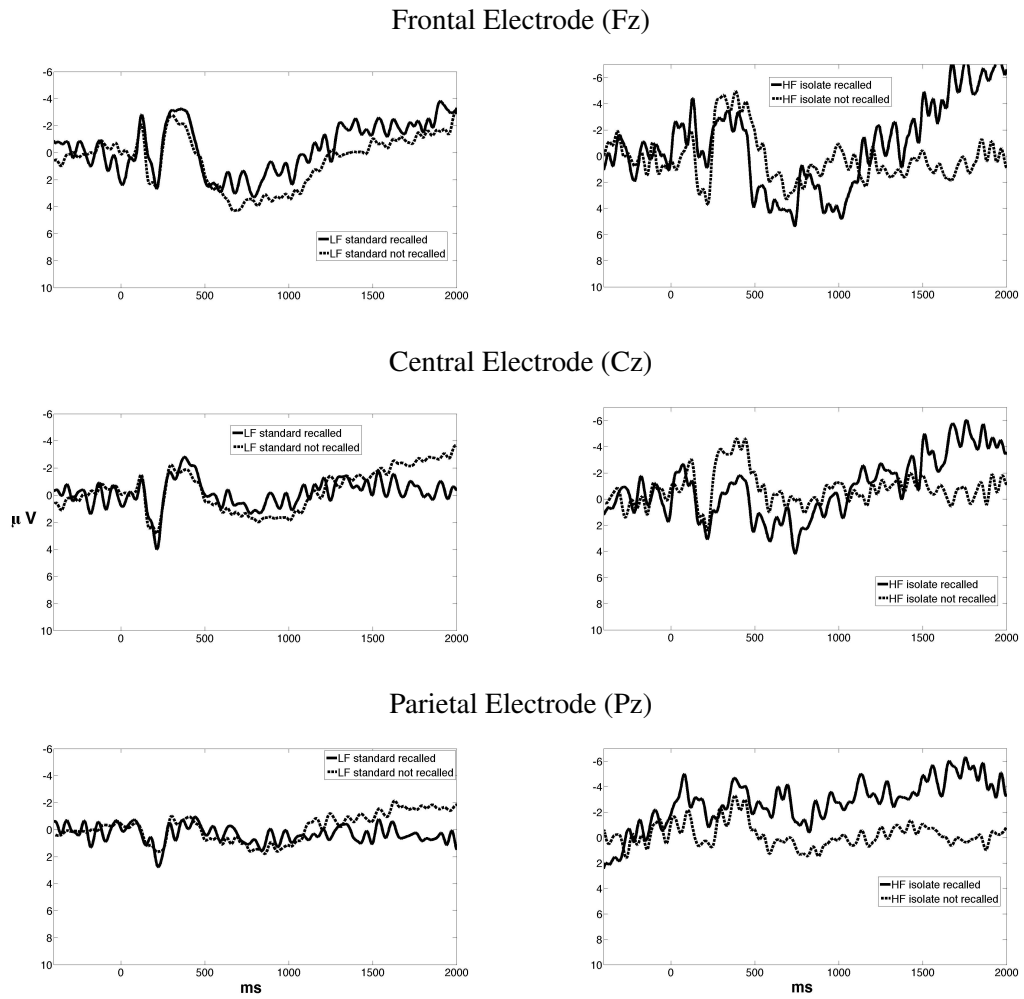


**Figure 4.:** Grand averages for Von Restorff lists. Grand average ERPs for three midline electrodes as a function of subsequent recall.

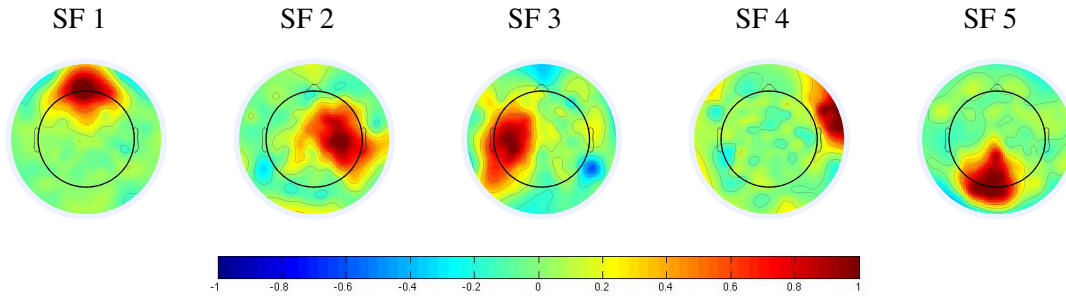


**Figure 5.:** Grand averages for HF lists including LF isolates. Grand average ERPs for three midline electrodes as a function of subsequent recall.





**Figure 6.:** Grand averages for LF lists including HF isolates. Grand average ERPs for three midline electrodes as a function of subsequent recall for LF lists including HF isolates.



**Figure 7.:** Spatial Factor Loadings. The loadings of the first five spatial factors obtained in the spatial PCA step.

right panel). It is conceivable that this component is an instance of the N400. To more accurately characterize and quantify the ERP components, the data were submitted to a spatio-temporal PCA.

### 3.3.1 PCA Results

A scree test on the PCA results indicated that 25 spatial factors and 8 temporal factors for each spatial factor should be retained and submitted to a Promax rotation. Note that a similar analysis using a Varimax rotation instead of Promax revealed spatial and temporal factors that were in close correspondence to the ones reported here. The reason that Promax was chosen as the preferred rotation for this study was that, possibly due to the large number of word types and conditions in this study (7 word types x 2 recall success categories = 14 conditions), the results of PROMAX appeared to be more clear-cut and easier to interpret. The results of the Varimax rotation were similar, but statistical comparisons were in many cases non-significant.

Spatial factor loadings of the first five spatial factors are shown in figure 7. Only spatial factors one, two, three and five were further analyzed, since the spatial distributions of factor four and all spatial factors beyond factor 5 did not appear to correspond meaningfully to ERP components. The total variance accounted for in the spatial step was 90.61%, and spatial factors one, two, three and five explained 21.24%, 8.99%, 8.97%, and 6.31% of the original variance, respectively.

The first three temporal factors of spatial factors 1, 2, 3 and 5 were analyzed in a 5 (word type: HF standard, LF standard, LF isolate, Size standard and Size isolate) by 2 (recall success) repeated measures ANOVA and significant interactions were followed up by subsequent planned comparisons. Furthermore, some additional planned comparisons were conducted when there were a-priori

reasons to believe that the ERPs might differ according to subsequent memory for a specific word type.

The HF isolates were analyzed in independent samples t-tests with uneven sample sizes, using the factor scores from 10 participants for the subsequently recalled, and from 13 participants for the not recalled HF isolates. Only significant results, results approaching significance, or results of specific interest with respect to our a priori hypotheses are reported here.

### 3.3.1.1 Frontal Factor

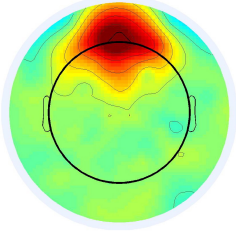
Spatial factor one exhibited a prefrontal distribution, as can be seen in figure 8 along with virtual ERPs and temporal factors. Due to the frontal distribution, a temporal factor most likely corresponding to the slow wave was of special interest, which appears to be captured in temporal factor 1 (figure 8 C). However, there were no indications for a main effect of word type, recall success, or an interaction for the PCA scores of this factor (all p-values bigger than .47).

Both inspection of the ERPs of frontal electrodes (e.g. Fz in figures 4, 5 and 6) and the virtual ERPs of the frontal factor (figure 8 B) show a strong positive-going potential peaking around 700 ms for all word types and both recalled and non-recalled words. For most word types, this positivity, which is captured in temporal factor 2 (see figure 8 C), did not appear to differ between subsequently recalled and subsequently not recalled words. However, the overall word type (5 levels: HF standard, LF standard, LF isolate, size standard, size isolate) by recall success interaction approached significance,  $F(4,48)=2.3, p=.07$ . A subsequent planned comparison indicated that for size isolates, the positivity was larger for subsequently recalled than not recalled words,  $t(12)=3.07, p<.01$ , while for LF standards the reverse was true, such that the amplitude was smaller for subsequently recalled compared to not recalled words,  $t(12)=2.4, p=.03$ . For the remaining three word types, there was no difference between recalled and not recalled words (all p-values bigger than .23).

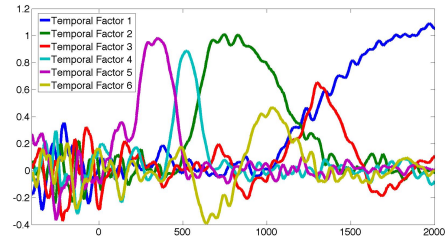
For the third temporal factor, peaking at about 1400 ms, the overall ANOVA resulted in a significant main effect for recall success,  $F(1,12)=6.67, p=.02$ . Thus, the factor scores for this factor were more positive for subsequently recalled compared to not recalled words across word types.

Note that none of the differences due to recall success for the HF isolates approached significance, indicating that any appearance of such differences in the virtual ERPs is likely due to the low number of trials and hence the high level of noise for this word type.

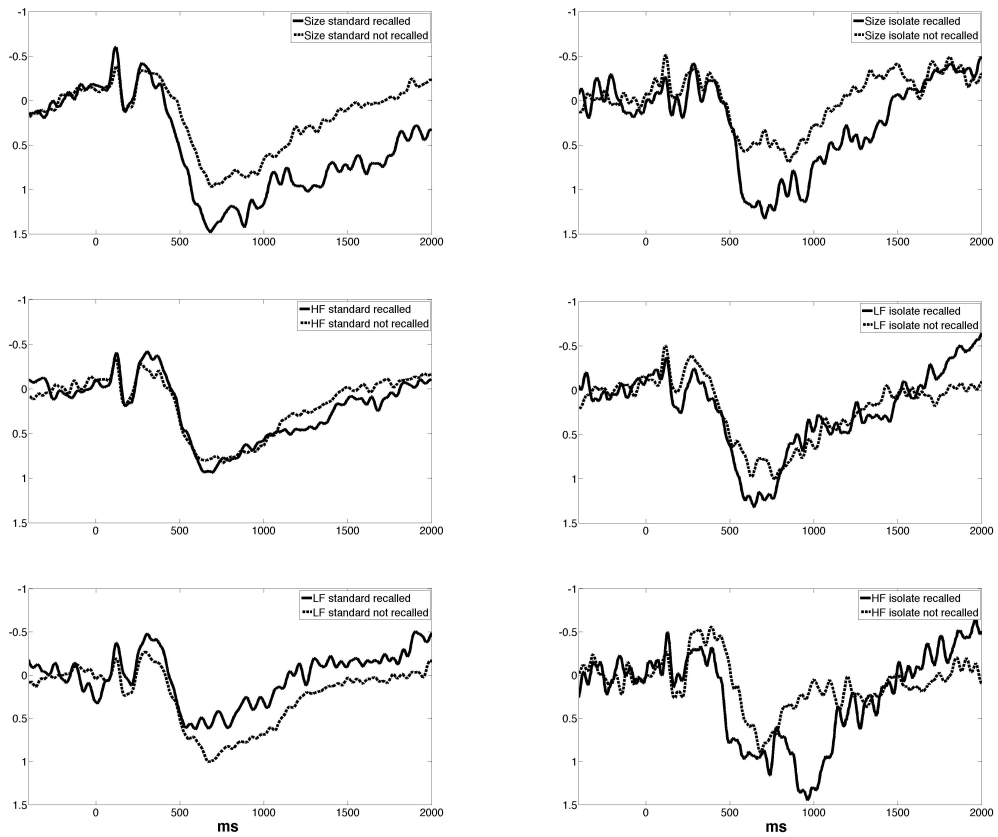
### A. Spatial Factor Loadings



### C. Temporal Factor Loadings



### B. Virtual ERPs



**Figure 8.:** Spatial factor 1. A. Spatial factor loadings. B. Virtual ERPs: Averaged spatial factor scores, plotted over time. C. Temporal factors loadings.

### 3.3.1.2 Right Lateralized Spatial Factor

Spatial factor two was right laterally distributed (figure 9). The negative peak around 400 ms in the virtual ERPs, possibly an instance of the N400, seemed to correlate with subsequent recall success for HF isolates (figure 9 B). However, neither visual inspection of grand average waveforms on right electrodes, nor inspection of the virtual ERPs for this spatial factor indicate any strong differences for subsequent recall success for any other word type.

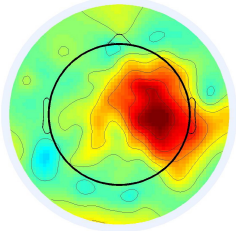
Statistical analysis of the first three temporal factors did not reveal any significant main effects or interactions involving the recall success comparison. However, the overall 5 (word type) by 2 (recall success) ANOVA on the factor scores for temporal factor 2 did result in a significant main effect for word type,  $F(4,48)=2.73$ ,  $p=.04$ . In a subsequent analysis collapsing across recall success, we found that the amplitude for this component was significantly more positive for size isolates than for all other word types (note, however, that the HF isolates were not included in this analysis). This temporal factor peaked around 600 ms (figure 9 C) and it is not obvious which previously reported ERP component it might correspond to.

Temporal factors 2 and 3 overlapped the time period in which ERPs appeared to differ between subsequently recalled and not recalled HF isolates. Indeed, temporal factor 2, peaking at 600 ms, was significantly more positive for recalled than not recalled words,  $t(21)=2.44$ ,  $p=.02$ . The difference for temporal factor 3, peaking at 400 ms and likely representing an instance of the N400, approached significance,  $t(21)=1.99$ ,  $p=.06$ . This indicates that the N400 factor tended to be more negative-going for not recalled than successfully recalled words. Note that although this N400 amplitude difference between HF isolates that were and that were not subsequently recalled approached significance, further planned comparisons revealed no overall differences in N400 amplitude between HF isolates and any other word type.

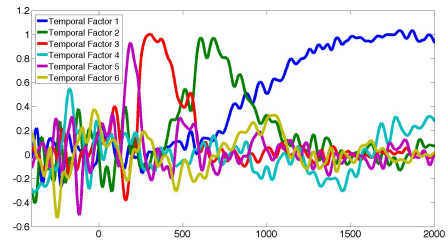
### 3.3.1.3 Left Lateralized Spatial Factor

The third spatial factor was left lateralized and its virtual ERPs appeared to morphologically resemble slow wave effects (see figure 10). Thus, the virtual ERPs exhibited a slow positive wave that appeared to differentiate between recalled and not recalled words for size isolates, LF isolates, and, in the reverse direction, LF standards (figure 10 B). The first and second temporal factor spanned the time period corresponding to this positive-going slow wave (figure 10 C). For temporal factor

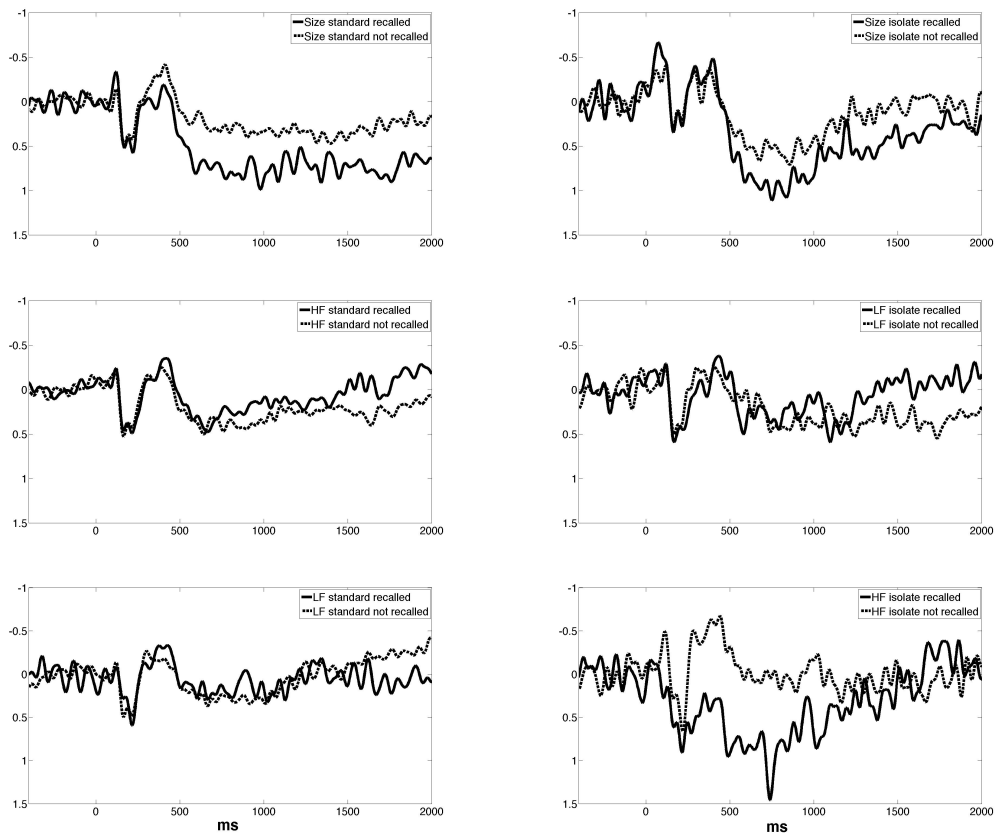
### A. Spatial Factor Loadings



### C. Temporal Factor Loadings

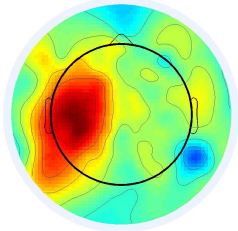


### B. Virtual ERPs

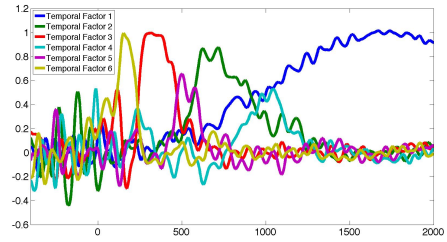


**Figure 9.:** Spatial factor 2. A. Spatial factor loadings. B. Virtual ERPs: Averaged spatial factor scores, plotted over time. C. Temporal factors loadings.

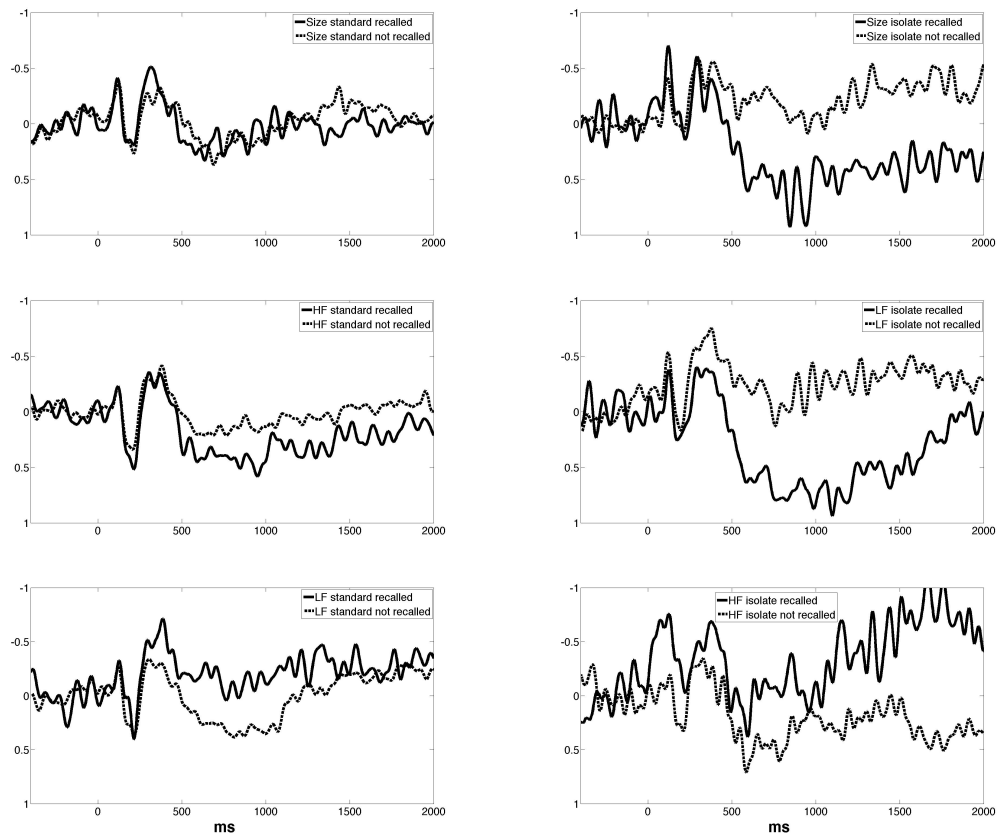
### A. Spatial Factor Loadings



### C. Temporal Factor Loadings



### B. Virtual ERPs



**Figure 10.:** Spatial factor 3. A. Spatial factor loadings. B. Virtual ERPs: Averaged spatial factor scores, plotted over time. C. Temporal factors loadings.

1 exhibiting a slowly increasing temporal distribution across the epoch, the overall 5 by 2 ANOVA did not reveal any significant effects or interactions. However, since the virtual ERPs strongly suggested differences due to recall for LF and size isolates, we performed further planned comparisons for only these two word types. For the size isolates, the difference between subsequently recalled and not recalled words was significant,  $t(12)=3.38$ ,  $p<.01$ , while for the LF isolates, the difference was non-significant,  $t(12)=1.5$ ,  $p=.16$ .

The second temporal factor overlapping the time period of the slow wave as visible in the virtual ERPs peaked at 650 ms (figure 10 B and C). The overall ANOVA on the scores of this factor revealed a significant word type by recall success interaction,  $F(4,48)=3.93$ ,  $p<.01$ . For LF isolates and size isolates, this factor was significantly more positive for recalled compared to not recalled words ( $t(12)=2.28$ ,  $p=.04$  and  $t(12)=3.14$ ,  $p<.01$ , respectively). Unexpectedly, for LF standards the relationship tended to be the opposite way. Thus, LF standards that were later not recalled tended to elicit a larger positivity than the ones that were later recalled,  $t(12)=1.88$ ,  $p=.08$ . Subsequent memory effects for the other word types did not approach significance.

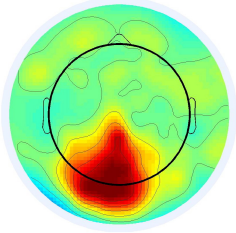
Note that although the virtual ERPs for HF isolates for spatial factor 3 appear to be different between subsequently recalled and subsequently not recalled words, none of the temporal factor scores obtained in the PCA differed significantly by recall success. Therefore, the visual impression of any difference due to recall success is likely due to a high level of noise caused by low trial numbers.

#### **3.3.1.4 Parietal Spatial Factor**

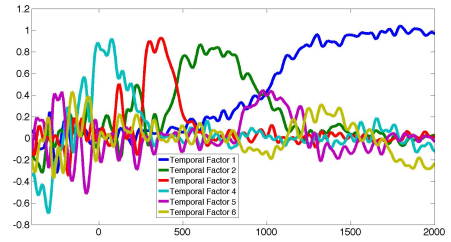
The spatial distribution of the fifth spatial factor is shown in figure 11, along with its virtual ERPs and temporal factor loadings. Spatial factor 5 showed a parietal distribution characteristic of the P300 (figure 11 A). Furthermore, a positive peak is present in the virtual ERPs for the size isolates but not size standards (figure 11 B), supporting the idea that this factor corresponds to the P300. Temporal factors 1 and 2 overlap with the time period of the P300 (figure 11 B and C). No significant effects or interactions were obtained for temporal factor 1, which apparently corresponds to a posterior slow wave. Temporal factor 2 seems to more closely resemble P300 morphology, although it peaks somewhat late, around 750 ms. For this temporal factor, the overall word type by recall success interaction approached significance,  $F(4,48)=2.39$ ,  $p=.06$ . Subsequent



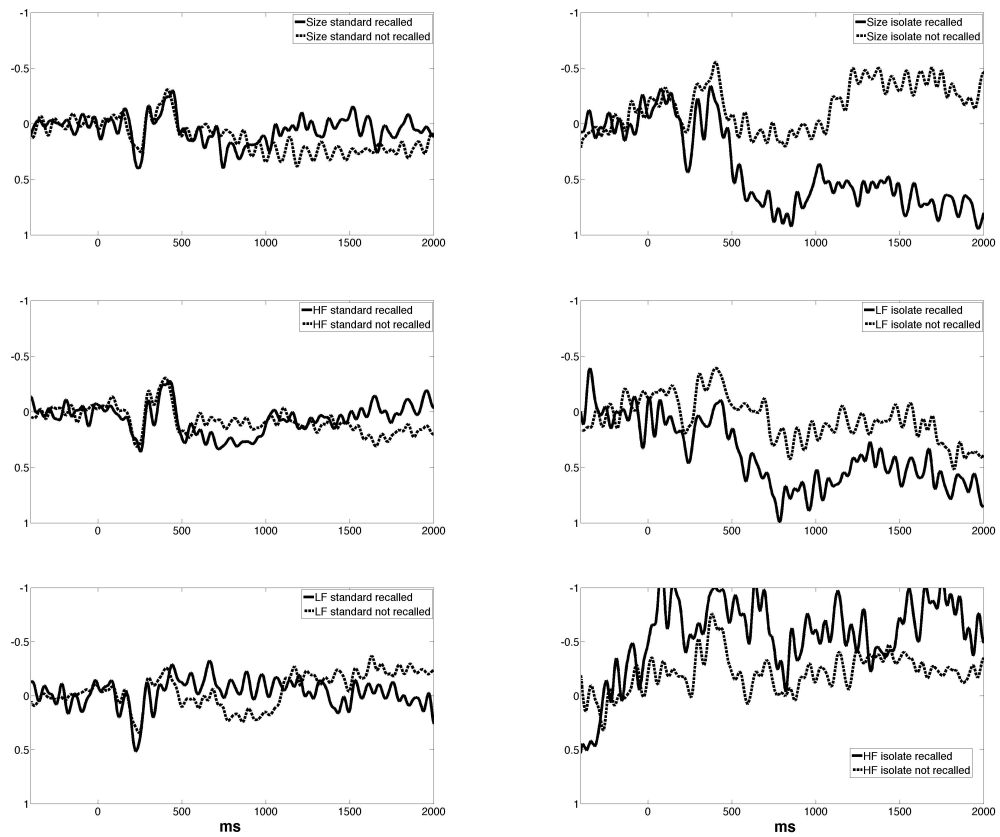
### A. Spatial Factor Loadings



### C. Temporal Factor Loadings



### B. Virtual ERPs



**Figure 11.:** Spatial factor 5. A. Spatial factor loadings. B. Virtual ERPs: Averaged spatial factor scores, plotted over time. C. Temporal factors loadings.

planned comparisons indicated that for LF isolates, factor scores were significantly more positive for subsequently recalled compared to not recalled words,  $t(12)=2.56$ ,  $p=.03$ . While the difference was in the same direction and appeared to be of about equal magnitude for size isolates, this comparison did not reach significance,  $t(12)=1.72$ ,  $p=0.11$ . For all other word types, there were no significant subsequent memory effects for the P300 factor.

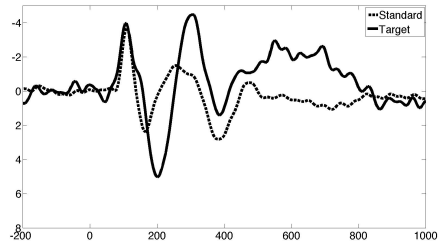
For temporal factor 3, showing a negative peak around 400 ms after the stimulus, the main effect for subsequent recall success approached significance,  $F(1,12)=3.23$ ,  $p=.1$ . This factor tended to be more negative for subsequently not recalled compared to recalled words, which is in line with the impression one might get from visually inspecting the virtual ERPs for size isolates and LF isolates. Note that although a similar effect seems to be present for the same factor elicited by the HF isolates, this difference was not significant. Furthermore, no other comparisons for subsequent recall for spatial factor 5 for HF isolates approached significance.

### **3.4 Oddball Task**

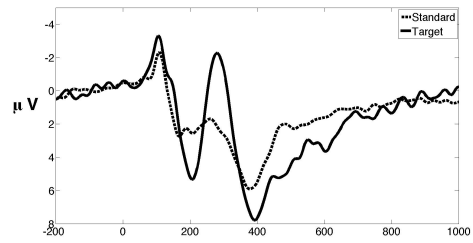
In the oddball task, all participants reached a response accuracy of at least 96.8% (mean: 99.5%) in responding to standards and at least 75% (mean: 90.7%) when responding to targets, indicating that every participant performed well above chance. The difference in response accuracy between standards and targets was significant,  $t(12)=4.23$ ,  $p<.01$ . Furthermore, reaction times were significantly longer to targets (mean: 451.6 ms) compared to standards (mean: 399.76 ms),  $t(12)=3.15$ ,  $p<.01$ . Figure 12 shows the grand average ERPs from three midline electrodes. A P300 is visible in the parietal electrode, although it seems as though the standards elicit a large positivity as well.

In the PCA, 8 spatial factors and 5 temporal factors were retained and submitted to a Promax rotation. By visual inspection, we determined that only spatial factors 1, 2, and 3 corresponded meaningfully to distributions of ERP components, so all further analysis focused on these three spatial factors. Spatial factors 1, 2 and 3 accounted for 38.65%, 25.55% and 9.5% of the original variance, respectively. Factor loadings, virtual ERPs and temporal factors from spatial factors 1-3 are shown in figure 13. Paired samples t-tests were conducted on the scores of the temporal factors for the first three spatial factors. Only statistically significant analyses at a level of .05 are reported here.

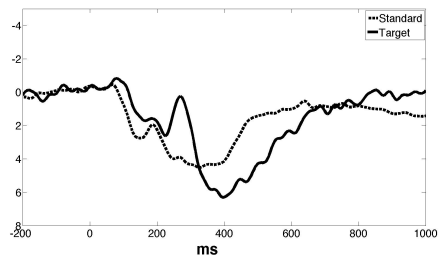
Frontal Electrode (Fz)



Central Electrode (Cz)



Parietal Electrode (Pz)

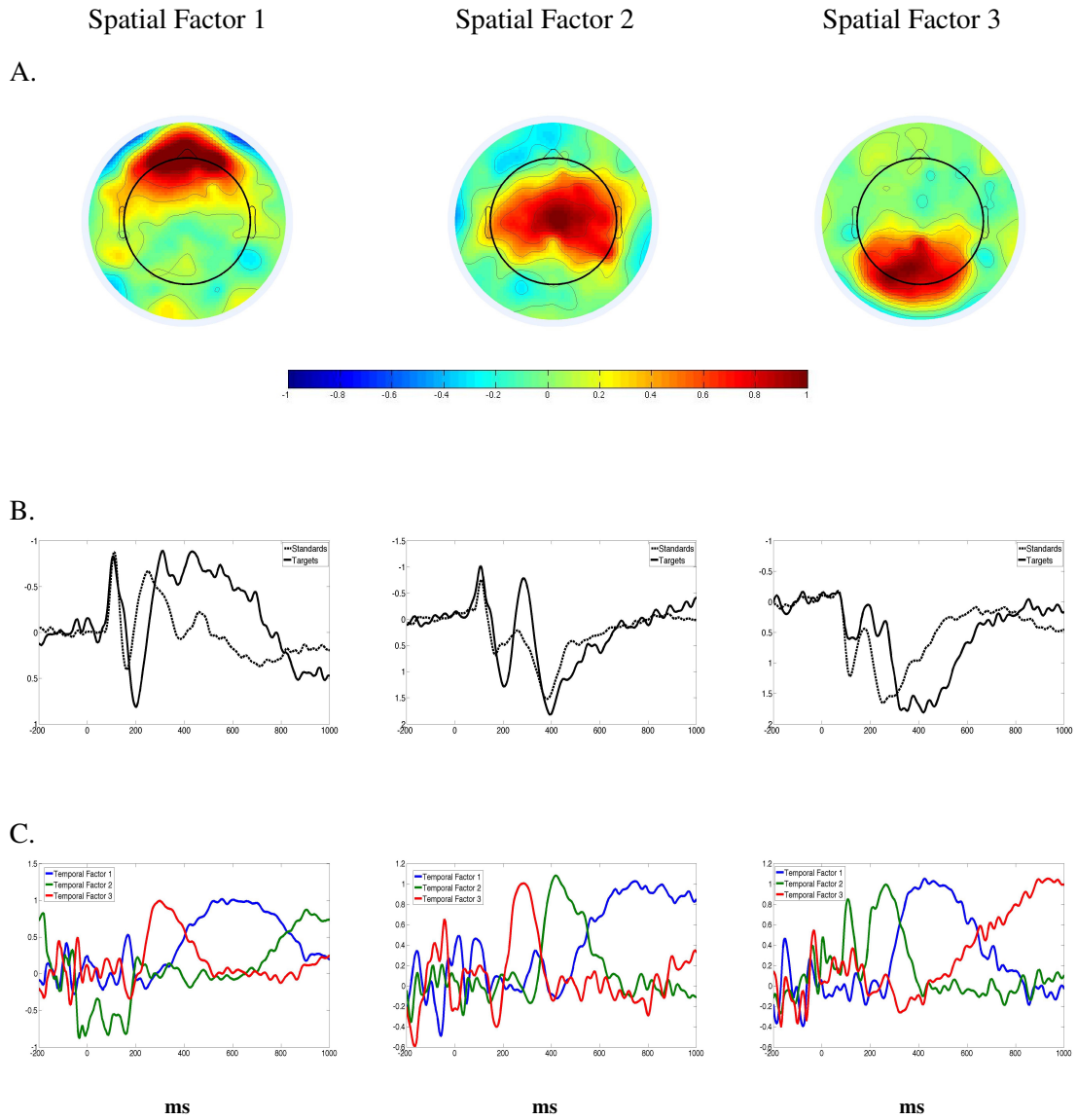


**Figure 12.:** Grand average ERPs from the oddball task.

As can be seen in figure 13 A and B, spatial factor 1 exhibited a frontal distribution and its virtual ERPs were characterized by a slow negative wave prominent for the target ERPs. The first temporal factor (figure 13 C), which appeared to capture this negative wave, was significantly larger for targets than standards,  $t(12)=2.61, p=.02$ . The second temporal factor exhibited an uninterpretable distribution, but differed between targets and standards,  $t(12)=2.33, p=.04$ . Furthermore, the fourth temporal factor (not shown in figure 13), capturing the large positivity peaking about 200 ms after the stimulus, was significantly more positive for targets than for standards,  $t(12)=4.17, p<.01$ .

Spatial factor 2 was centrally distributed and showed a positive peak around 200 ms, followed by a negative peak around 300 ms, with both peaks apparently more pronounced for targets. Statistical analysis of the temporal factor corresponding to the positive peak (temporal factor 5, not shown in figure 13,  $t(12)=2.37, p=.04$ ), as well as the temporal factor corresponding to the negative peak (temporal factor 3, see figure 13 C,  $t(12)=2.42, p=.03$ ) confirmed this impression.

The third spatial factor was parietally distributed, a distribution that is characteristic of the P300. Note that the spatial distribution of this factor was remarkably similar to the distribution of the P300 factor obtained from the ERPs in the memory task. One difficulty that we encountered in the statistical analysis of the factor scores of the P300 component in the oddball task was that the standards seemed to elicit a large P300 as well. However, this positivity elicited by the standards peaked at a shorter latency, as visible in the virtual ERPs as well as the raw averaged ERPs (see figure 12, bottom panel, and figure 13 B, right panel). Thus, the temporal factor corresponding to the P300 elicited by the targets (temporal factor 1, figure 13 C) did not seem to be the same temporal factor that corresponds to the parietal positivity elicited by the standards (temporal factor 2). Therefore, in the first step we compared the factor scores for temporal factors 1 and 2 between the targets and standards, and in the second step we compared the factor scores for temporal factor 1 for the targets with the factor scores for temporal factor 2 for standards, a comparison that seems to be more valid in reflecting differences in P300 amplitude. Factor scores of temporal factor 1 were significantly larger for target ERPs compared to standard ERPs,  $t(12)=2.99, p=.01$ , while factor scores for temporal factor 2 were larger for the standards,  $t(12)=4.13, p<.01$ . However, factor scores for temporal factor 1 for the targets did not significantly differ from factor scores of temporal factor 2 for the standards,  $t(12)=0.01, p=.99$ . None of the remaining temporal factors for spatial factor 5 differed between targets and standards.



**Figure 13:** The first three spatial factors for the oddball task. A. Spatial factor loadings; B. Virtual ERPs; C. Temporal factor loadings for the first three temporal factors.

## Chapter 4

### Discussion

#### 4.1 Summary of Results

Behaviorally, we obtained a classical Von Restorff effect for physically isolated words, a typical word frequency effect for non-isolated words, as well as a reversal of the word frequency effect when HF and LF words occurred isolated in lists of opposite-frequency words. The PCA on the subject ERPs revealed a frontal factor, a right lateralized factor possibly in part corresponding to the N400, a left lateralized factor that morphologically resembled the frontal slow wave, as well as a parietally distributed factor representing the P300.

In line with our expectations, the size isolates elicited a P300 component peaking around 750 ms after the stimulus. The difference in P300 amplitude between subsequently recalled and not recalled size isolates, as measured by the corresponding PCA factor scores, approached significance. Size isolates also elicited a left lateralized slow wave whose amplitude showed a clear and significant subsequent memory effect. Another positive component was present in the frontal spatial factor, peaked around 700 ms after the stimulus and was larger for subsequently recalled compared to not recalled size isolates. Lastly, a positivity in the right lateralized spatial factor, also peaking 700 ms after the stimulus, was significantly larger for size isolates than for any other word type, but did not differ with respect to subsequent recall.

Similar to the size isolates, LF isolates also elicited a P300 that was significantly more positive for recalled compared to not recalled words. Furthermore, the left lateralized slow wave also showed a subsequent memory effect such that later recalled LF isolates elicited more positive amplitudes than not recalled LF isolates. This difference, however, was only statistically significant for the earlier temporal factor overlapping the slow wave. The frontal and right lateralized spatial factors did not contain any temporal factors with significant subsequent memory effects for LF isolates.

The ERP components and subsequent memory effects we obtained for HF isolates were remarkably different from those for LF isolates. Thus, HF isolates did not appear to elicit a strong P300 or a left lateralized slow wave component, nor were any subsequent memory effects significant for these factors. Instead, the right lateralized negativity peaking at 400 ms after the stimulus, which is most likely an instance of the N400, was negatively correlated with subsequent recall. That is, HF isolates that were not recalled elicited a larger N400 than later recalled words. Furthermore, the positivity following the N400 was larger for subsequently recalled compared to not recalled HF isolates.

For non-isolated LF words (LF standards), we also obtained a couple of interesting findings. First, there was a significant subsequent memory effect for the frontal positivity peaking at 700 ms, such that subsequently recalled LF standards elicited smaller positive amplitudes than LF standards that were later not recalled. Note that this subsequent memory effect is in the opposite direction as the one obtained for size isolates. Secondly, the left lateralized slow wave component tended to be more positive for later recalled compared to later not recalled LF standards. Again, this subsequent memory effect is in the reverse direction than for size isolates and LF isolates.

The HF standards and size standards did not elicit any subsequent memory effects unique to their word type. However, there were two main effects for subsequent recall success that did not depend on word type. First, in the frontal factor a positivity peaking at 1400 ms after the stimulus differed between subsequently recalled and not recalled words. It is not obvious which ERP component this effect corresponds to, but it seems that the effect represents the sustained positivity that is apparent in the virtual ERPs for subsequently recalled size standards, size isolates and HF standards (see figure 8 B). Secondly, the overall subsequent memory effect in the amplitudes of the parietal negative peak at 400 ms approached significance.

For the oddball task data, the PCA resulted in three interpretable spatial factors. The first one was a frontal factor characterized by a narrow peak at 200 ms followed by a slow negative potential. Both of the corresponding temporal factors differed significantly between targets and standards. Secondly, a central factor showed a positive peak at 200 ms, followed by a negative peak at 300 ms and another positive peak at 400 ms. The first positivity as well as the negativity differed between targets and standards. Lastly, and of biggest interest for the purpose of the current study, the third spatial factor had a parietal distribution and a strong positive-going virtual ERP that most likely

corresponds to the P300. The spatial distribution of this factor was in close correspondence to the P300 obtained in the memory experiment. Unexpectedly, the P300 latency differed between targets and standards, and when the appropriate temporal factor scores were analyzed, there was no significant amplitude difference between the two stimulus types.

Our main findings will now be discussed in more detail in the light of previous studies and theoretical frameworks.

#### **4.2 Correlations of ERP Amplitudes and Recall in the Von Restorff Paradigm**

Our behavioral and ERP data from the Von Restorff lists replicate and extend findings of previous studies on ERP correlates of the Von Restorff paradigm. Size isolates were more likely to be recalled, resulting in a strong and significant Von Restorff effect. Furthermore, words that were displayed in a larger font than the other words elicited a P300. The amplitude of the P300 elicited by size isolates was correlated with subsequent recall, such that later recalled words elicited larger P300 amplitudes than later not recalled words (although this effect only approached significance).

The P300 factor obtained in the PCA showed a posterior distribution that was in close correspondence to the P300 elicited in the standard oddball paradigm. The latency of the P300 in the current study was surprisingly long - it peaked between 700 and 800 ms after the deviant stimulus. In previous, analogous Von Restorff memory experiments, P300 latency ranged between 500 and 600 ms (Fabiani & Donchin, 1995; Otten & Donchin, 2000; Karis et al., 1984). P300 latency is thought to reflect the time it takes the subject to evaluate the distinctive property of a stimulus (Kutas et al., 1977). Since the distinctiveness attribute in the Von Restorff lists was a physical manipulation of font size, however, it should have been detected and evaluated relatively quickly. Furthermore, since the manipulation was analogous to other studies, the expected latency should have been similar to these studies. Therefore, the reason for the long P300 latency in this study remains to be elucidated.

The correlation between P300 amplitude and subsequent recall success under rote memory instructions replicates a large number of previous studies (Fabiani et al., 1986; Fabiani, Karis, & Donchin, 1990; Fabiani & Donchin, 1995; Otten & Donchin, 2000). However, in the current study this subsequent memory effect only approached statistical significance for size isolates. The non-significant finding could have resulted from low statistical power, or from large between subject variability, for example due to the use of different memorization strategies.



Since we replicated previous findings on correlations between P300 amplitude and subsequent recall when words are physically distinctive and rote memorization is used, our data are in line with the previous theoretical framework used to explain this finding. The idea is that P300 amplitude correlates with subsequent memory when a distinctive attribute is tagged onto the memory trace and at test used for successful retrieval. The relationship between P300 and recall only holds true when participants use an item-based memorization strategy. If the participant engages in elaboration of item-to-item relationships at study, the outcomes of this elaboration become more effective for retrieval so that retrieval does not necessarily rely on the distinctiveness attribute (Fabiani & Donchin, 1995; Donchin & Fabiani, 1991).

In our study, size isolates also elicited a left lateralized slow wave component that was strongly and significantly correlated with subsequent recall. In previous studies, a frontally distributed slow wave was correlated with subsequent recall for non-isolated words and for isolates only when elaborative strategies were used (Fabiani, Karis, & Donchin, 1990), but there was usually no such correlation when participants used rote memory (Otten & Donchin, 2000). Similarly, Karis, Fabiani and Donchin found that while the slow wave was correlated with recall across all memorization strategies, the effect was *strongest* when elaborative strategies were used (Karis et al., 1984). Our findings are somewhat in contrast with these studies, because we found a very pronounced correlation between slow wave amplitudes and recall although participants were instructed to use rote memory, an instruction that was followed by most participants according to self-reports.

The slow wave in the early studies was reported to have a frontal maximum (Fabiani & Donchin, 1995; Otten & Donchin, 2000), although later cortical imaging of the subsequent memory ERP effects in the Von Restorff paradigm mapped them in part to the *left* prefrontal cortex (Ding et al., 2002; Lian et al., 2002). The left lateralization of the slow wave factor in our study is consistent with these cortical imaging studies. However, the spatial factor loadings in our study were more posterior (i.e. less frontal) than expected. One could suggest that the rotation method used in the PCA skewed the actual spatial distribution of the component in the raw data. However, it should be noted that the use of a different rotation, like Varimax, resulted in a very similar distribution of this slow wave factor, indicating that the PCA rotation method is unlikely to be the only explanation for our finding. Therefore, it remains unclear whether the left lateralized slow wave obtained in the current study corresponds to the frontal slow wave reported in previous studies.

In conclusion, in our study both a positive slow wave and the P300 were correlated with subsequent recall for size isolates under the same memorization method, which is somewhat in contrast with previous studies reporting a dissociation between correlations of the two components' correlations to recall depending on word type and memorization strategy. It is possible that participants used a mixture of memorization strategies (without admitting this during debriefing), resulting in the correlation of both the P300 and the frontal slow wave to subsequent recall in our study.

The componential investigation of our findings with a spatio-temporal PCA allowed us to discover two other positivities elicited by size isolates that temporally overlapped with the P300 but exhibited different spatial distributions. Note that in most previous studies, smaller numbers of electrodes limited the possibilities to separate apart components with similar temporal, but different spatial distributions. Therefore, the findings discussed next are novel. The first one of these positive components was present in the frontal factor and its amplitude was significantly correlated with subsequent recall for size isolates. Although its distribution may be slightly more frontal than would be expected based on previous PCA studies (Dien et al., 2003), this factor may represent an instance of the P3a (N. K. Squires et al., 1975), or novelty P3 (Courchesne, Hillyard, & Galambos, 1975). In the first study on the P3a, this component was elicited by infrequent stimuli in unattended oddball sequences (N. K. Squires et al., 1975). The novelty P3 is a component with a comparable spatio-temporal distribution to the P3a and is elicited in modified oddball sequences by a third type of stimulus, which is novel and task irrelevant (Courchesne et al., 1975; Dien et al., 2003). While some authors refer to the P3a and the P300 as two "sub-components" of the same ERP component (Polich, Howard, & Starr, 1983), they have different spatial distributions and respond differently to experimental manipulations and should therefore be regarded as two separate components.

Based on previous knowledge about the eliciting conditions of the P3a (or the novelty P3), the elicitation of this component in the current study is not entirely surprising. Thus, while obviously each word in a study - test memory paradigm is relevant for the subject's task, the larger font size of the isolate *per se* is not task relevant. It is therefore well conceivable that the frontal positivity correlated with subsequent recall for size isolates in our paradigm is an instance of the P3a. The correlation of this component to subsequent recall is a novel finding and needs further exploration in future studies. For example, it remains to be elucidated whether the presence of this correlation depends on memorization strategy.

Another additional positivity elicited by size isolates, which also peaked at 700 ms and therefore overlapped with the P300, was found in the right lateralized spatial factor. This positivity was larger for size isolates than for any other word type and could also possibly be classified as a P3a. However, this component was not correlated with subsequent recall.

To portray a complete picture of ERP correlates to subsequent recall in the Von Restorff paradigm, it is important to also consider subsequent memory effects for size standards (i.e. the regular-sized words that were studied Von Restorff lists). Previous studies have indicated that for these non-isolated words, the frontal slow wave is correlated with subsequent recall (Fabiani, Karis, & Donchin, 1990). In the current study we were unable to replicate this finding. That is, we did not obtain any correlations between ERP amplitudes elicited by size standards and subsequent recall except for the overall effects for subsequent recall across word types, that will be discussed later. Most importantly, the left lateralized slow wave that showed strong correlations with subsequent recall for size isolates (and for LF isolates, as will be discussed later) was not significantly correlated with recall for size standards. The discrepancy between our results and previous findings is unclear, since our experimental paradigm closely resembled previous studies.

### **4.3 Word Frequency Effects and Their Reversal due to List Composition**

The word frequency effect on recall memory for pure lists refers to the superior recall of HF compared to LF words (Mandler et al., 1982). In our study, due to the presence of only one word of deviating frequency, the LF and HF lists could be considered close-to-pure, resulting in the prediction that this word frequency effect on recall for pure lists should be observed for LF and HF standards. Indeed, the HF standards were significantly better recalled than LF standards at primacy and plateau positions.

An additional pattern of interest is that word frequency effects on recall depend on list composition, such that when mixed lists are studied, LF words are recalled equally well or better than HF words (Watkins et al., 2000; Fernández et al., 2002; Hulme et al., 2003; McDaniel & Geraci, 2006). In the current study, we investigated possible explanations for this abolition of the word frequency effect in mixed lists by inserting an individual word with a frequency that was opposite of the rest of the list (LF and HF isolates). We expected that these isolated words would behave similar to words in mixed lists of LF and HF words (although quantitative differences are likely, because in

most previous studies mixed lists consisted of 50% LF and 50% HF words, as opposed to 14 words of one type and only one word of the other type in our study).

We found that the isolation in lists of opposite-frequency words increased the probability of recall for LF words, but decreased recall levels for HF words. That is, the probability of recall was higher for LF isolates than for LF standards, but lower for HF isolates than for HF standards. This finding is in line with the prediction that LF words lead to equal or superior recall than HF words when studied in mixed lists, which is the typical reversal of the word frequency effect on recall in mixed lists.

It is important to note that each LF word and each HF word was equally likely to be selected as an isolate, because isolates and standards were selected at random from the same pools of HF and LF words. Furthermore, the words that some participants encountered as LF and HF isolates were encountered by other participants as LF and HF standards. Therefore, differences in linguistic properties between isolates and standards are not likely to explain the enhancement of recall for LF words or reduction of recall for HF words due to their isolation.

It is worth noting that our behavioral findings on word frequency effects on recall are different from the Von Restorff effect. Von Restorff emphasized the importance of comparing each isolated stimulus to recall levels of non-isolated items in the same list (that were hence of a different stimulus type), as well as to recall levels of the same item type when the stimulus did not occur in isolation (Von Restorff, 1933). She found enhanced recall for the isolated item compared to *both* non-isolated items within the same sequence and non-isolated items of the same stimulus type when occurring as a non-isolate in a different list. In our study, isolated LF words were recalled at an *equally high* level of HF standards at comparable serial positions, while isolated HF words were recalled an *equally low* level as LF standards. Therefore, the LF isolates were better recalled than non-isolates of the same type in different lists (LF standards), but not compared to the non-isolated words in the same list (HF standards).

While LF isolates did show one characteristic of the Von Restorff effect (enhanced recall compared to non-isolated words of the same type), HF isolates did not show any indications for a memory enhancing isolation effect. Thus, they were recalled at an equal level as non-isolated words in the same list (LF standards), but at an even *lower* level than non-isolated words of the same type (HF standards).

How do our behavioral data on the reversal of the word frequency effect on recall compare to other studies that manipulated list composition beyond comparing pure lists with mixed lists of 50% HF and 50% LF words? In a serial recall experiment, DeLosh and McDaniel (1996) manipulated list composition such that lists contained (1) 8 LF words, (2) 8 HF words, (3) 2 LF and 6 HF words, or (4) 2 HF and 6 LF words. Recall was better for pure HF lists than for any other list, but recall was on a comparable level for list types 2, 3 and 4. While this study's paradigm was quite different from ours, these results are consistent with the absence of a Von Restorff effect in our study when words were isolated due to their word frequency. Another study by Saint-Aubin and LeBlanc investigated effects of list composition on serial recall (Saint-Aubin & LeBlanc, 2005). These authors had participants study either pure lists or lists of five words of the same frequency and one word of the opposite frequency. In addition to the usual word frequency effect for pure lists, they found that an isolated LF word in a list of HF words was equally likely to be recalled as the HF words in the same list. While the experimental paradigm was again very different from ours, the findings are consistent with the lack of a Von Restorff effect for LF isolates in our study. However, in contrast to our findings, these authors also found an advantage for the isolated HF word compared to the LF words in the same list. This finding is inconsistent with our result that HF isolates lead to equally low recall levels as LF standards.

Our behavioral findings appear to suggest a major role for inter-item elaborative processes in the reversal of the frequency effect. According to this idea, LF words' recall is elevated to the same level of HF words' recall when they occur in lists of HF words, because the latter facilitate elaborative processes between list items. For the same reason, HF words' recall is reduced to the level of LF word's recall when these occur isolated in LF lists, because the LF lists do not facilitate elaborative processes between list items to an equal extent. Thus, the theory that the outcomes of item-to-item associative processes are used as retrieval aids for recall, and that the level of such associations depends on the ratio of HF and LF words in a list (Hulme et al., 2003) seems to be a likely candidate to explain the current study's behavioral findings.

Are influences of distinctiveness on recall also a possible explanation for our behavioral results? As described above, if the HF and LF isolates were perceived as distinctive compared to the rest of the list in a similar way as a large word in a list of regular-sized words, they would be expected to produce *higher* level of recall than the non-isolated words in the same list. We did not obtain

such an effect in the current study. However, there is one possibility that may uphold an influence of distinctiveness at encoding time on recall: That there is some intrinsic property of HF words that makes them more likely to be recalled than LF words. Isolated LF words, however, are perceived as distinctive in HF lists leading to an isolation effect. The general memory advantage for HF words and the distinctiveness advantage for the LF isolate then elevate retrieval to an equally high level, resulting in the apparent absence of a Von Restorff effect observed in our data. It is unclear, however, how this possible explanation would account for the impairment in recall of HF words when they occur in isolation. This idea will be further discussed after the outline of ERP correlates associated with word frequency effects on recall. Overall, if we only consider the behavioral results, they appear to favor theories based on elaborative processes over distinctiveness-based explanations of the word frequency effect's reversal in mixed lists.

As noted in the introduction, a specific pattern of behavioral findings can be associated with different ERP effects (Otten & Donchin, 2000). Therefore, in order to test the distinctiveness and inter-item elaboration theories for the reversal of the word frequency effect in mixed lists, we investigated which ERP components were pronounced and correlated with subsequent recall for LF and HF standards and isolates. Specifically, these ERP effects were compared to the subsequent memory effects found in the Von Restorff lists discussed above to investigate whether subjective distinctiveness plays a role in word frequency effects on recall as well. These data are discussed next.

#### **4.4 Common Processes of the Von Restorff Effect and Word Frequency Effects?**

The current study was designed to investigate whether word frequency effects on recall, specifically the reversal of the word frequency effect in mixed lists, share underlying cognitive processes with the Von Restorff effect in recall. We approached this question by comparing the ERPs and their subsequent memory effects elicited by physically deviant words to those elicited by words of different frequencies when they occurred clustered or in isolation. First, the discussion in this section will focus on ERPs elicited by the LF isolates. We directly compare the LF isolates to size isolates, and it seemed that the two word types elicited very similar ERP components and subsequent memory effects. Furthermore, we will compare the analysis of output order for LF isolates and size isolates during the recall test to investigate whether both isolation types are available as

organizational aids for retrieval. Afterwards, we will discuss the subsequent memory effects for HF and LF standards, followed by a section on the quite different subsequent memory effects elicited by HF isolates. Next follows a section on the two subsequent memory effects that were found across all word types (including size isolates and standards). The conclusion of this section will focus on the overall implications of our findings for theories on the reversal of the word frequency effect in recall.

#### **4.4.1 ERP Components and Correlations to Recall for LF Isolates**

The LF isolates represent the word type that is most critical for testing our hypotheses, because these are the words that could be either perceived as distinctive, or associated with stronger inter-item elaborative processes than LF standards. To distinguish between these possibilities is the main purpose for the current study.

Our data indicated that LF words isolated in lists of HF words were associated with very similar ERP signatures and subsequent memory effects compared to size isolates. First, LF isolates elicited a P300 component whose amplitude was strongly and significantly correlated to subsequent recall. This finding seems to indicate that when LF words occur in lists of HF words, they are indeed perceived as distinctive and that this distinctiveness attribute of the memory trace is used effectively as a retrieval aid. This, in turn, indicates that the reversal of the word frequency effect in mixed lists can in part be explained by enhanced subjective distinctiveness of LF words when HF words are present. Therefore, the ERP data of our study imply that common processes partly underlie the effects of physical distinctiveness, and word frequency on recall.

This implication somewhat conflicts with the conclusions one may draw from the behavioral results, as discussed in the previous section. Thus, if LF isolates are indeed perceived as distinctive, and if this distinctiveness attribute of the memory trace is effectively used in retrieval processes as implied by the correlation of P300 amplitude with recall, one could argue that recall should be enhanced for LF isolates compared to *both* LF standards and HF standards. In contrast to this expectation, the recall of LF isolates was on an equal level as for HF standards at the corresponding serial positions. There are two possible explanations for this discrepancy between our behavioral and ERP findings.

First, as outlined in the previous section, LF words may be inherently more difficult to recall than HF words. Retrieval processes may be more effective for HF words in general because their easier lexical access may facilitate the reconstruction of the item from an incomplete or noisy memory trace - a process that has been named “redintegration” (Hulme et al., 1997). At the same time, recall of LF words isolated in HF lists benefits from their subjective distinctiveness in a similar way as physically distinctive words, elevating recall levels for LF isolates as well. The enhancing effects of facilitated redintegration processes for HF words and the enhancing effect of distinctiveness for LF isolates then elevate recall to comparable levels. A problem with this explanation is that it attributes the general superiority of HF words in recall to item-based effects at retrieval without attributing any importance on study context. This is inconsistent with the reduced probability of recall for HF isolates compared to HF standards. That is, if recall of HF words is enhanced because memory traces of HF words are easier to reconstruct at retrieval, this should be true for all HF words regardless of study list composition. It appears, therefore, that the enhancement of recall levels of HF standards and LF isolates to the same level due to facilitated redintegration processes and distinctiveness effects, respectively, is an unlikely explanation for our findings.

Another possible explanation for the discrepant findings between behavioral and ERP results is that due to variance in certain characteristics, some LF words stand out from lists of HF words while other LF words do not stand out to an equal extent. In other words, some LF isolates may be perceived as distinctive, elicit a large P300 and are successfully recalled. Other LF isolates, in turn, are not perceived as distinctive, that is, they “blend in” with the HF words, they do not elicit much of a P300, and do not lead to enhanced recall. This mixture of LF isolates perceived as distinctive and LF isolates not perceived as distinctive compared to the HF standards, then, led to the correlation between P300 and subsequent recall as well as attenuated levels of recall for LF isolates. In line with this idea, it is apparent from the outline of different properties that may make LF words distinctive (see introduction), that there is a high variability in characteristics within the group of LF words.

Which characteristics could cause only a subset of LF isolates to be subjectively perceived as distinctive when occurring in lists of HF words? Given that word frequency is usually considered a measure of prior exposure to a word, it appears that the individual’s subjective familiarity with each word may play a role. The idea is that if a LF isolate is very unfamiliar to the participant, it is likely to stand out from a list of HF words, while LF isolates that are rather familiar to the participant



may be perceived as similar to the rest of the list. If this is the explanation for our findings, LF isolates that were later recalled should be associated with lower subjective familiarity ratings than LF isolates that were later not recalled. An analysis of the subjective familiarity ratings given in the questionnaire at the end of the experiment, however, did not support this hypothesis. Thus, subsequently recalled LF isolates and subsequently not recalled LF isolates received about equal familiarity ratings that were not significantly different (see figure 2 B).

There are a number of other properties of LF isolates that may have led to a subjective distinctiveness of only a subset of the LF isolates. As outlined in the introduction, LF words tend to differ from HF words in terms of orthographic or phonological uniqueness (Criss & Malmberg, 2008; Landauer & Streeter, 1973), low levels of context variability (Adelman et al., 2006; Steyvers & Malmberg, 2003), or low numbers of connections to other words in lexical memory (Cofer & Shevitz, 1952; Nelson & McEvoy, 2000). Different levels one or more of these characteristics for different LF isolates may have led to the subjective perception of only some LF isolates as distinctive. Our current analysis does not allow for any strong conclusions about which characteristics may have played a crucial role.

Regardless of which explanation accounts for the discrepancy between our ERP findings and our behavioral findings, LF words in lists of HF words did elicit a P300 which was correlated with subsequent recall. This is the ERP pattern obtained for size isolates in our study and many previous studies when rote memory was used. Conclusively, outcomes of context updating processes elicited by physically deviant and by LF words in lists of HF words appear to be effective retrieval cues at time of recall. The reversal of word frequency effects on recall in mixed lists may therefore be explained in part by distinctiveness effects comparable to those elicited in the Von Restorff paradigm.

In addition to the P300, LF isolates also elicited a left lateralized positive slow wave - the same component we also found for size isolates in the Von Restorff lists. Like for size isolates, this component was correlated with subsequent recall for LF isolates. The difference between recalled and not recalled words was, however, only statistically significant during the earlier temporal factor overlapping the slow wave effect (about 700 to 1000 ms after the stimulus) for LF isolates.

As noted above, it is unclear whether this left lateralized slow wave component reflects the frontal positive slow wave reported in previous studies, which has been associated with inter-item elabo-

rative memorization processes. If future research reveals that the two components are the same, our findings can be considered evidence for the involvement of enhanced elaborative processes associated with LF words occurring in HF word lists compared to both HF and LF standards. An additional explanation for the reversal of word frequency effects on recall in mixed lists is therefore that outcomes of item-to-item elaboration LF words embedded in lists of HF words are enhanced and effectively used for retrieval.

The frontal positivity that was correlated with subsequent recall for size isolates, did not show such a correlation for LF isolates. As we suggested above, this positivity may be an instance of the P3a. It appears, therefore, that the amplitude of this component is positively associated with subsequent recall for physical deviant stimuli, but not for deviance in terms of word frequency.

In summary, LF isolates were associated with ERP subsequent memory effects that showed striking similarities to size isolates. They elicited a P300 that was correlated with subsequent recall. Furthermore, the left lateralized slow wave was also correlated to subsequent recall. These findings indicate that both subjective distinctiveness at study time as well as working memory processes supporting elaborative processes between items are the basis for enhanced recall of LF words in lists of HF words. In other words, both item-specific and elaborative processes appear to play a role in the reversal of word frequency effects on recall.

#### **4.4.1.1 Comparison of Output Order for LF Isolates and Size Isolates**

In a previously published ERP study on the Von Restorff paradigm, Fabiani and Donchin investigated the effect of isolation type - comparing semantic and physical isolates - on the probability that this isolate will be written down first or last on the recall sheet (Fabiani & Donchin, 1995). The idea was that if the isolation feature is available as an organizational aid, the isolated item should be retrieved somewhat separately from the other items in the list. The authors found that physical isolates were most likely to be recalled last while semantic isolates were more likely to be recalled first, compared to all other output positions. The authors interpreted this finding such that both semantic and physical isolation features are available as organizational aids in memory retrieval, but that the two isolation features are used differently from each other.

In a similar comparison, we found that both size isolates and LF isolates tended to be written down at the end of the recall sheet. However, this pattern was only significant for LF isolates. This

finding once more corroborates that the distinctiveness of LF isolates and physical isolates act upon memory processes in a similar way (at least when the results for physical isolates of Fabiani and Donchin's study are taken into account as well). Thus, both LF isolates and size isolates appear to be similarly available as organizational aids in memory.

Note that an alternative interpretation of our results on output order is that they are only a byproduct of the specific retrieval strategy. Thus, participants may first recall the last words of the study lists, then continue with words from the beginning of the lists, and finally proceed to the middle list positions. Since isolates always occurred at serial positions 6-10, these words would automatically tend to be written down last during recall. Our current analysis cannot rule out this possibility, so further analysis is needed to draw more definite conclusions.

#### **4.4.2 ERP Components and Correlations to Recall for HF and LF Standards**

In order to study underlying neuro-cognitive processes of word frequency effects on memory, not only ERP components elicited by the words isolated in lists of opposite-frequency words, but also those elicited by the HF and LF standards are of interest. Comparable to the non-isolated words in the Von Restorff lists, the HF standards did not elicit significant subsequent memory effects for any spatio-temporal factors that were unique to this word type. The only two overall subsequent memory effects that appeared to be present across word types will be discussed in a later section of this chapter.

LF standards, in turn, elicited two distinct subsequent memory effects which are worth discussing. First, the frontal positivity that was positively correlated with subsequent recall for size isolates, was negatively correlated with recall for LF standards. That is, LF standards that were later recalled elicited smaller frontal positivities than LF standards that were later not recalled. As discussed above, this ERP component may be an instance of the P3a. If the component in our study indeed reflects a P3a, this finding could indicate that physical deviance that may be inherent to some LF words (possibly due to unusual letter combinations) actually *harms* their probability of subsequent recall when they occur clustered with other LF words. A possible interpretation is that due to larger difficulty for inter-item elaborative processes between LF words, those words that are somewhat deviant as reflected in the P3a are integrated with the other words even more inefficiently. For this reason, the presence of a large P3a may harm recall for LF standards. This explanation is rather

speculative, and it is unclear why some LF standards may elicit a larger P3a but show no evidence of a P300. Therefore, more research is needed to clarify the negative correlation of the amplitude of our frontal positive component to subsequent recall for LF standards.

The second component that tended to differ between recalled and not recalled LF standards was the left lateralized positive slow wave. Again, the correlation of this component's amplitude with recall success was in the opposite direction than for size isolates. That is, subsequently recalled LF standards elicited *smaller* slow wave amplitudes than subsequently not recalled LF standards. This finding is again difficult to interpret - if this slow wave component indeed plays a role in inter-item elaborative encoding processes, one would expect that the LF standards that elicit larger component amplitudes will be more likely to be recalled. It is unclear why stronger elaborative processes should lead to worse recall for LF standards and therefore the explanation for our finding needs to be clarified in further research.

Conclusively, our overall findings on subsequent memory effects of LF standards do not appear to have obvious interpretations. It is clear, however, that both subsequent memory effects for LF standards are in the opposite direction than the subsequent memory effects for the same components elicited by size isolates. This suggests that different neuro-cognitive processes at encoding are effective for later retrieval of LF words that occur in clusters compared to physically deviant words. It is therefore not likely that word frequency effects on recall when pure lists are studied can be attributed in any extent to processes that also lead to effects of subjective distinctiveness on recall as observed in the Von Restorff paradigm.

#### **4.4.3 Correlations of N400 Amplitudes and Recall Success for HF Isolates**

The HF isolates are the last word type that needs discussion to get a complete picture of ERP components and their correlations to subsequent recall for word frequency effects in mixed lists. As outlined above, the HF isolates are an important control to the LF isolates. Thus, any effects that occur for *both* HF and LF isolates would need to be interpreted such that their cause is not the frequency of the word *per se*, but rather whether it *deviates* from the frequency of the surrounding words. If the same effects we found for LF isolates are not obtained for HF isolates, interpretations can incorporate ideas of secondary distinctiveness accounts that imply that LF words, but not HF words exhibit some level of distinctiveness.

It is clear in our data, that ERP subsequent memory effects were qualitatively very different for HF isolates compared to LF isolates. Therefore, the correlations to subsequent recall for P300 and slow wave appear to be specific to LF isolates and size isolates, and are not simply due to the fact that the frequency of the isolate differs from the rest of the list. Hence, it seems justified to conclude that LF isolates in lists of HF words, but not HF isolates in lists of LF words are subjectively perceived as distinctive similarly to words that deviate due to font size.

Despite the lack of P300 and slow wave correlations to recall for HF isolates, we found a correlation between N400 amplitude and subsequent recall. Thus, HF isolates that were later recalled elicited a smaller negative N400 amplitude than words that were later not recalled. Furthermore, the positivity following the N400 with the same spatial distribution was larger for subsequently recalled HF isolates than for not recalled HF isolates. This finding was unexpected, since previous studies did not find correlations between N400 amplitude and subsequent recall (Fabiani & Donchin, 1995). Note that in our study, overall N400 amplitude did not differ between HF isolates compared to any other word type.

As outlined in the introduction, the N400 has been implicated in semantic integration processes (Kutas & Hillyard, 1980). Its amplitude is larger for unexpected words that are in a different semantic category than the expected word compared to unexpected words that are within the expected category (Federmeier & Kutas, 1999). This and other findings imply that the more effort is needed to successfully integrate a word into the semantic context, the larger is the N400 amplitude. The results of the current study indicated that it is more likely for a HF isolate to be recalled when it elicits a small N400 amplitude. Combined with previous knowledge on the N400, our findings suggest that the less effort is needed to semantically integrate a HF word into the semantic context established by a list of LF words, the more likely it is that this word will later be recalled. Semantic deviation, which has been suggested as a different type of distinctiveness (Fabiani, 2006), is therefore not beneficial for encoding and retrieving HF words isolated in lists of LF words.

It is unclear how this finding can be interpreted with respect to word frequency effects on recall. Generally, there should not be semantic differences between our pool of HF and LF words. However, if there were any semantic differences, an N400 effect should also be obtained for LF isolates, which we did not observe in our study. Therefore, the correlation between N400 amplitude and subsequent recall for HF isolates is probably based on natural trial-by-trial variability in N400

amplitudes. Why N400 related processes at encoding could harm subsequent recall for HF isolates remains an open question.

Lastly, it is important to note that due to the small trial numbers of HF isolates, the ERP data were quite noisy. It is possible that some additional subsequent effects for HF isolates were obscured by these high levels of noise in the data. However, visual inspection of the ERPs do not indicate that there is even a tendency for a P300 or a slow wave effect for HF isolates, so it appears safe to conclude that these were specific to LF and size isolates.

#### **4.4.4 Subsequent Memory Effects Across Word Types**

There were two main effects for subsequent recall success that seemed to be present across word types. Both of these effects were unexpected and appear to raise more questions than they answer. However, since these subsequent memory effects were found for all word types it is worth to briefly discuss each of them. It is important to keep in mind that due to three missing data points, the HF isolates were not included in the overall statistical test that resulted in these main effects, so the following findings may not generalize to this word type.

The first overall subsequent memory effect was obtained for a positivity that peaked at 1400 ms after the stimulus, which was present in the frontal factor. This factor appears to represent the sustained positivity that is apparent in the virtual ERPs for subsequently recalled size standards, size isolates and HF standards (see figure 8 B). It may be related to frontal slow wave effects, but shows an unusual temporal distribution. Therefore, it is unclear which ERP component this factor may correspond to and how to interpret the functional significance of its correlation to subsequent recall.

Secondly, the overall subsequent memory effect in the amplitude of the parietal negative peak at 400 ms approached significance. Thus, this component was larger for subsequently not recalled words compared to recalled words. It remains to be investigated whether there is a possible relationship between this component and the right lateralized N400 component.

#### **4.4.5 Implications of our Data for Word Frequency Effects on Recall**

Overall, our data provide strong support for the distinctiveness hypothesis of the reversal of word frequency effects on recall when mixed lists are studied. Although the behavioral data on LF isolates

do not exactly resemble the Von Restorff effect, a P300 was elicited and correlated with subsequent recall for LF words that occurred as isolates in lists of HF words. This is the same ERP effect as observed for size isolates in our study as well as several previous studies. This finding supports the idea that subjects perceive LF words as distinctive when they are surrounded by HF words, and that this perceived distinctiveness aids recall in a similar way as a physical distinctiveness attribute. HF isolates, in turn, do not appear to bear subjective distinctiveness, indicating that the low word frequency is necessary for the perceived distinctiveness, and not only whether the frequency of a word *differs* from the rest of the list. Conclusively, the memory enhancing effect of LF words' perceived distinctiveness when they are surrounded by HF words may play a large role in the reversal of word frequency effects on recall in mixed lists.

An open question is what characteristic of the LF words make them subjectively distinctive compared to the HF words, eliciting the P300 and enhancing the probability of recall. The distinctive attribute does not appear to be the subjective familiarity, since we found no difference in familiarity ratings for recalled and not recalled LF isolates. One possibility is that unusual letter combinations in the LF words make these words distinctive compared to other words (Criss & Malmberg, 2008). This "orthographic distinctiveness" explanation would imply that perceptual, or physical, deviance plays a role in our findings, comparable to the Von Restorff effect.

A second question that requires further investigation is how the probability of a LF word occurring in a list affects whether it is perceived as distinctive as indexed by the P300 and whether this distinctiveness attribute is effectively used at retrieval. Thus, we have repeatedly stated that our findings may provide explanations for reversal of word frequency effects in mixed lists. However, typically such findings are obtained in lists containing 50% of each frequency category. In our study, only one of 15 words was of opposite frequency. There is strong prior evidence that the higher the probability is that a deviant event occurs, the smaller is the amplitude of the P300 (Duncan-Johnson & Donchin, 1977), and thereby the event's perceived distinctiveness. Therefore, it remains to be investigated whether LF words still elicit a P300 that is correlated to recall when the probability of a LF to occur on a given trial is closer to 50%. In other words, before our data can be generalized to other list compositions, lists containing a larger proportion of LF words will need to be studied in a similar paradigm to ours.

As discussed before, our data also provide support for an inter-item elaboration theory on the word frequency effect on recall as well as its reversal in mixed lists. It is likely that the left slow wave represents the frontal slow wave found in previous studies. This component was pronounced and correlated to recall for both LF isolates and size isolates. It appears, therefore, that item-to-item elaborative processes at encoding are enhanced for LF words embedded in HF lists, and that these are used as effective retrieval cues at the time of the memory test. One problem with this interpretation is that HF standards did not elicit such a slow wave and a correlation of its amplitude to subsequent recall. This issue will be discussed in more detail in the next section.

Note that our data cannot distinguish between different types of elaborative processes that may be the basis for the reversal of the word frequency effect. Not enough is known about the slow wave to draw any conclusions about which types of elaborative processes it is associated with. Therefore, whether the relevant elaborative processes are the encoding of order information (DeLosh & McDaniel, 1996; Merritt et al., 2006), the activation of pre-existing associations (Hulme et al., 2003), or yet other types of inter-item encoding, remains an open question.

In conclusion, our data are in line with theories suggesting that both item-based processes as well as inter-item elaborative processes need to be considered to explain the reversed word frequency effect on recall memory when mixed lists of both HF and LF words are studied (DeLosh & McDaniel, 1996; Tse & Altarriba, 2007).

#### **4.5 Implications of the Current Study for the Frontal Slow Wave**

In addition to the implications for processes underlying word frequency effects on recall, our data also provide new information on the conditions under which the slow wave is pronounced and thereby for interpretations of its functional significance. As discussed above, one new finding worth discussing is that our spatio-temporal PCA revealed a different spatial distribution of the slow wave component than expected based on previous findings. Thus, the slow wave previously reported showed a frontal maximum, but our slow wave was left lateralized. While it remains unclear whether the slow wave found in our study corresponds to the slow wave observed in previous studies, its temporal morphology suggests so.

If our slow wave indeed is an instance of the “frontal positive slow wave” found in previous studies on a similar paradigm (Fabiani, Karis, & Donchin, 1990; Karis et al., 1984; Otten & Donchin,



2000), our findings raise some additional questions. Thus, pure lists of HF words are thought to facilitate item-to-item elaborative processes, while pure lists of LF words do so to a lesser extent (see previous section). If we assume that HF words elicit higher levels of inter-item elaborative processes, and if we further assume that the slow wave manifests such elaborative processes, then the prediction would be that HF words elicit larger slow wave amplitudes than LF words in pure lists. For example, previous findings by Fabiani, Karis and Donchin (1990) indicated that when subjects used elaborative strategies, the frontal slow wave was more pronounced than when subjects used rote memorization. However, in our study there was no significant difference between the slow wave amplitude elicited by HF standards and LF standards.

Furthermore, if for HF words, outcomes of item-to-item associative processes are more likely to be used as retrieval aids than for LF words, we should also observe a stronger *correlation* between slow wave amplitudes and subsequent recall for HF standards than for LF standards. Again, we did not observe this pattern of results. In the contrary, there was no significant correlation between slow wave amplitudes and recall for HF standards. For LF standards, there was even a tendency for a negative correlation between slow wave amplitudes and recall, but only for the earlier temporal factor overlapping the slow wave.

So, if HF lists facilitate item-to-item processing, then why did HF standards not elicit larger slow wave amplitudes and correlations to recall than LF standards? One important point to note is that we manipulated characteristics of the stimulus material, not which encoding processes the participant would actively and consciously engage in to memorize it. Fabiani, Karis and Donchin (1995) had participants study lists in which some words were semantically related using two different item-based processing tasks - either a semantic task or a structural task. One may predict that in such a design, the semantic task would direct the participant to process the semantic relatedness between stimuli more intensively than with the structural task and that due to these enhanced inter-item processes, the frontal slow wave should be more pronounced for the semantic task. However, there were no indications in their data to support this prediction. An important parallel between the study by Fabiani and Donchin (1995) and the current study is that both studies did not manipulate whether or not the participant was actively engaged in elaborative processing at study time. In their study, the semantic judgement was made for each item individually, and in our study, participants were instructed to use rote memory. Therefore, in both studies, if stronger item-to-item elaborative

processes took place in one of the conditions, these were a by-product of the stimulus material or the processing task, respectively, and may therefore have been automatic and uncontrolled.

In conclusion, our data contribute to previous knowledge on the slow wave indicating that it is pronounced and correlated with subsequent recall for non-isolated words only when participants are engaged in a task that requires active, controlled processing of stimulus material in working memory - a process that likely relies on the prefrontal cortex.

It remains to be investigated whether instructions to use item-to-item elaborative strategies, such as forming sentences with the study words (as opposed to rote memorization instructions like in the current experiment), have different effects on slow waves elicited by HF words when compared to LF words. That is, if these inter-item elaborative tasks are easier to accomplish for lists of HF words than for LF words, and if these tasks are furthermore associated with increased slow wave amplitudes, one may find a difference in slow wave amplitudes and their correlation to subsequent recall for pure HF lists compared to LF lists when such processing tasks are used.

#### **4.6 Abnormalities in the Current Study's Oddball Paradigm**

In standard oddball paradigms, the infrequent, task relevant stimuli elicit a P300, but the frequent stimuli do not elicit a P300 (Donchin, 1981; Duncan-Johnson & Donchin, 1977; Kutas et al., 1977). The oddball paradigm in our study was typical in that the "O" occurred with a low probability (20%), while the "X"s occurred with a high probability (80%). Thus, in this paradigm we expected to find a large P300 amplitude in the average ERPs elicited by the infrequent "O"s, but not for the frequent "X"s. In contrast to this expectation, we found large parietal positivities for both the infrequent *and* the frequent stimuli, while the latency was shorter for the frequent. When the appropriate temporal factors were statistically compared, there was no significant difference between the P300 amplitudes elicited by frequent and infrequent.

This finding was unexpected and the reason for this unusual finding is unclear. It is unlikely that the finding is due to an abnormal sample, since an analogous finding was obtained with the identical paradigm used on a different sample of about 60 students (Ty Brumback, personal communication). One potential influence on these unusual results may stem from the requirement of each participant to respond to each stimulus, including frequent *and* infrequent. However, previous studies found that independently of which stimulus the subject responds to, a task-relevant infrequent stimulus

will elicit a P300 while a task-relevant frequent stimulus will not (K. C. Squires, Donchin, Herning, & McCarthy, 1977). Therefore, requiring participants to respond to both stimuli is an unlikely explanation the absence of amplitude differences in the P300 elicited by standards and targets in our study.

Another possible explanation is that each trial ended at the time the participant gave their response. Therefore, the inter-trial interval (ISI) for any two trials depended on the reaction time to the first of the two stimuli. While the influence of the ISI on P300 amplitude has been studied between blocks of oddball stimuli (Polich, 1990), to our knowledge there have not been any studies on variable ISI's *within* blocks.

A further possibility is that the perceptual salience of our stimuli caused a P300 to be elicited by every stimulus, including standard trials. Thus, both the “X” and the “O” occupied a large part of the screen and were therefore possibly larger than in most previous studies on the oddball paradigm. Perceptual salience does seem to play a role in auditory oddball paradigms, such that the P300 is largest for tones that are task-relevant, loud and infrequent (K. C. Squires et al., 1977).

Finally, error rates and reaction times to targets were significantly larger than for standards, suggesting that there was a response bias toward responses to the standard stimuli. This response bias may be another possible influence on our unusual findings.

Therefore, further analysis will need to focus on why the standards in our oddball paradigm appear to elicit a P300 that is of shorter latency than the P300 for targets. It is possible that only a subset of the standards elicited a P300, driving the appearance of a P300 in the overall ERPs. This possibility is subject to further investigation.

## Chapter 5

### Conclusions and Future Directions

Overall, our study supports the idea that LF words are recalled equally well or better than HF words when they occur in mixed lists because they are subjectively perceived as distinctive in these types of lists. Similarly to the processes suggested for physical isolates, this distinctiveness attribute is tagged to the memory trace and can be used as an effective retrieval cue. This perceived distinctiveness in mixed lists is specific to LF words, and does not occur for HF words isolated in lists of LF words. Conclusively, similar neuro-cognitive processes appear to support the Von Restorff effect as well as the reversal of the word frequency effect on recall in mixed lists.

It stands to reason why our behavioral results did not reveal a Von Restorff effect for LF isolates although the correlations between P300 amplitude and subsequent recall supported the notion that distinctiveness aided recall for LF isolates. It is likely that a paradigm using a recognition memory test will reveal better memory for LF isolates than HF standards. Whether analogous ERP subsequent memory effects for the P300 will be observed for the LF isolates in a recognition paradigm is subject to further research.

Our study also supports an involvement of enhanced inter-item elaborative processes for LF words in lists of HF words as compared to LF words in clusters as a factor in the reversal of the word frequency effect on recall. Thus, correlations between ERP amplitudes and recall were not only found for the P300, but also for the positive slow wave, an ERP component that has been implicated in elaborative processes. It appears, therefore, that a combination of *both* item-specific distinctiveness processes as well as between-item elaboration causes the reversal of word frequency effects on recall in mixed lists.

Future studies are needed to investigate whether our conclusions are generalizable to mixed lists composed of different proportions of HF and LF words than used in the current study. Specifically, an important question is whether the same pattern of results would be obtained in a mixed list with 50% HF and 50% LF words, since this is the list composition in which the reversal of the

word frequency effect due to mixed lists is usually studied behaviorally. Another open question is how the participant's processing task at encoding affects the results. Correlations between P300 amplitude and recall for physical isolates have been reported only when rote memory is used. It is possible, therefore, that the correlation of P300 amplitude to recall for LF isolates depends on the memorization strategy in a similar way.

While our findings provide some new information on the eliciting conditions of the slow wave, the data also raised some new questions. That is, why is the slow wave left lateralized in our study, but showed a frontal maximum in previous studies? The question whether the component in our study the same as in previous studies will need further investigation. Another issue that needs clarification in further studies is why HF standards did not elicit larger slow wave amplitudes than LF standards although they are thought to facilitate inter-item associative processes.

It is unclear whether the correlations between the frontal positivity and subsequent recall for size isolates in our study actually reflect the P3a. To clarify whether P3a-related processes may aid successful memory encoding, future studies could investigate this question in a three-stimulus oddball paradigm followed by an incidental memory test for the novel items. It would be a great contribution to the current knowledge about the functional significance of the P3a to determine whether it is correlated to subsequent memory success in such a design.

Finally, further analysis is necessary to clarify the reasons for our unusual findings in the oddball task. That is, did only a subset of the frequent stimuli elicit a P300 like ERP component? If so, what are the characteristics of this subset of frequent stimuli, and why did they elicit a P300? A single-trial analysis identifying which ones of the frequent stimuli may have elicited a P300 will help clarify these issues.

## References

- Adelman, J. S., Brown, G. D. A., & Quesada, J. F. (2006). Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychological Science, 34*(3), 495-513.
- Allen, M., Badecker, W., & Osterhout, L. (2003). Morphological analysis in sentence processing: An erp study. *Language and Cognitive Processes, 18*(4), 405-430.
- Azizian, A., & Polich, J. (2007). Evidence for attentional gradient in the serial position memory curve from event-related potentials. *Journal of Cognitive Neuroscience, 19*(12), 2071-2081.
- Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory and Cognition, 29*(4), 639-647.
- Barber, H., Vergara, M., & Carreiras, M. (2004). Syllable-frequency effects in visual word recognition: evidence from erps. *NeuroReport, 15*(3), 545-548.
- Bosch, V., Mecklinger, A., & Friederici, A. D. (2001). Slow cortical potentials during retention of object, spatial and verbal information. *Cognitive Brain Research, 10*, 219-237.
- Chee, M. W. L., Westphal, C., Goh, J., Graham, S., & Song, A. W. (2003). Word frequency and subsequent memory effects studied using event-related fmri. *NeuroImage, 20*, 1042-1051.
- Cofer, C. N., & Shevitz, R. (1952). Word-associations as a function of word-frequency. *American Journal of Psychology, 65*(1), 75-79.
- Connine, C. M., Mullennix, J., Shernoff, E., & Yelen, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*(6), 1084-1096.
- Courchesne, E., Hillyard, S. A., & Galambos, R. (1975). Stimulus novelty, task relevance and the visual evoked potential in man. *Electroencephalography and clinical Neurophysiology, 39*, 131-143.

- Criss, A. H., & Malmberg, K. J. (2008). Evidence in favor of the early-phase elevated-attention hypothesis: The effects of letter frequency and object frequency. *Journal of Memory and Language*, *59*, 331-345.
- Davidson, D. (2006). Distinctiveness and memory. In R. R. Hunt & J. B. Worthen (Eds.), (p. 157-179). New York: Oxford University Press.
- Deldin, P., Deveney, C., Kim, A., Casas, B., & Best, J. (2001). A slow wave investigation of working memory biases in mood disorders. *Journal of Abnormal Psychology*, *110*(2), 267-281.
- Delorme, A., & Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics. *Journal of Neuroscience Methods*, *134*, 9-21.
- DeLosh, E. L., & McDaniel, M. A. (1996). The role of order information in free recall: Application to the word-frequency effect. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *22*(5), 1136-1146.
- Diana, R. A., & Reder, L. M. (2006). The low-frequency encoding disadvantage: Word frequency affects processing demands. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *32*(4), 805-815.
- Diedrich, O., Naumann, E., Maier, S., & Becker, G. (1997). A frontal positive slow wave in the erp associated with emotional slides. *Journal of Psychophysiology*, *11*(1), 71-84.
- Dien, J. (2010). The erp pca toolkit: An open source program for advanced statistical analysis of event-related potential data. *Journal of Neuroscience Methods*, *187*(1), 138-145.
- Dien, J., Spencer, K. M., & Donchin, E. (2003). Localization of the event-related potential novelty response as defined by principal component analysis. *Cognitive Brain Research*, *17*(3), 637-650.
- Ding, L., Goldstein, A., Lian, J., Donchin, E., & He, B. (2002). Cortical imaging of p300 component elicited by isolate words using different memory rehearsal strategies. *International Journal of Bioelectromagnetism*, *4*(2), 269-270.
- Donchin, E. (1981). Surprise! ... surprise? *Psychophysiology*, *18*(5), 493-513.
- Donchin, E., & Fabiani, M. (1991). Handbook of cognitive psychophysiology. central and autonomic nervous system approaches. In J. Jennings & M. Coles (Eds.), (p. 472-510). John Wiley and Sons.

- Duncan-Johnson, C. C., & Donchin, E. (1977). On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology*, *14*(5), 456-467.
- Fabiani, M. (2006). Distinctiveness and memory. In I. R. R. Hunt & J. B. Worthen (Eds.), (p. 339-360). New York: Oxford University Press.
- Fabiani, M., & Donchin, E. (1995). Encoding processes and memory organization: A model of the von restorff effect. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *21*(1), 224-240.
- Fabiani, M., Gratton, G., Chiarenza, G. A., & Donchin, E. (1990). A psychophysiological investigation of the von restorff paradigm in children. *Journal of Psychophysiology*, *4*, 15-24.
- Fabiani, M., Karis, D., & Donchin, E. (1986). P300 and recall in an incidental memory paradigm. *Psychophysiology*, *23*(3), 298-308.
- Fabiani, M., Karis, D., & Donchin, E. (1990). Effects of mnemonic strategy manipulation in a von restorff paradigm. *Electroencephalography and clinical Neurophysiology*, *75*, 22-35.
- Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, *41*, 469-495.
- Fernández, G., Klaver, P., Fell, J., Grunwald, T., & Elger, C. E. (2002). Human declarative memory formation: Segregating rhinal and hippocampal contributions. *Hippocampus*, *12*(514-519).
- Fernandez, G., Weyerts, H., Tendolkar, I., Smid, H. G. O. M., Scholz, M., & Heinze, H. (1998). Event-related potentials of verbal encoding into episodic memory: Dissociation between the effects of subsequent memory performance and distinctiveness. *Psychophysiology*, *35*(06), 709-720.
- Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming times. *Journal of Verbal Learning and Verbal Behavior*, *12*, 627-635.
- Francis, W., & Kucera, H. (1982). *Frequency analysis of english usage: Lexicon and grammar*. Boston: Houghton Mifflin.
- Gerhand, S., & Barry, C. (1999). Age-of-acquisition and frequency effects in speeded word naming. *Cognition*, *73*, B27-B36.
- Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory and Cognition*, *13*(1), 8-20.



- Glanzer, M., & Adams, J. K. (1990). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 16(1), 5-16.
- Glanzer, M., & Bowles, N. (1976). Analysis of the word-frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory*, 2(1), 21-31.
- Guo, C., Zhu, Y., Ding, J., Fan, S., & Paller, K. A. (2004). An electrophysiological investigation of memory encoding, depth of processing, and word frequency in humans. *Neuroscience Letters*, 356, 79-82.
- Hulme, C., Roodenrys, S., Schweickert, R., Brown, G., Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(5), 1217-1232.
- Hulme, C., Stuart, G., Brown, G. D. A., & Morin, C. (2003). High- and low-frequency words are recalled equally well in alternating lists: Evidence for associative effects in serial recall. *Journal of Memory and Language*, 49, 500-518.
- Humphreys, M. S. (1976). Relational information and the context effect in recognition memory. *Memory and Cognition*, 4(2), 221-232.
- Hunt, R. R. (2006). Distinctiveness and memory. In R. R. Hunt & J. B. Worthen (Eds.), (p. 3-25). New York: Oxford University Press.
- Hunt, R. R., & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior*, 20, 497-514.
- Karis, D., Fabiani, M., & Donchin, E. (1984). "p300" and memory: Individual differences in the von restorff effect. *Cognitive Psychology*, 16, 177-216.
- Kim, A. S. N., Vallesi, A., Picton, T. W., & Tulving, E. (2009). Cognitive association formation in episodic memory: Evidence from event-related potentials. *Neuropsychologia*, 47, 3162-3173.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203-205.
- Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: the p300 as a measure of stimulus evaluation time. *Science*, 197(4305), 792-795.

- Kutas, M., Van Petten, C., & Kluender, R. (2007). Handbook of psycholinguistics. In M. Gernsbacher & M. Traxler (Eds.), (2nd ed., pp. 659–724). Elsevier.
- Landauer, T. K., & Streeter, L. A. (1973). Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Verbal Learning and Verbal Behavior*, *12*, 119-131.
- Lian, J., Goldstein, A., Donchin, E., & He, B. (2002). Cortical potential imaging of episodic memory encoding. *Brain Topography*, *15*(1), 29-36.
- Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT press.
- Malmberg, K. J., & Murnane, K. (2002). List composition and the word-frequency effect for recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(4), 616-630.
- Malmberg, K. J., & Nelson, T. O. (2003). The word-frequency effect for recognition memory and the elevated-attention hypothesis. *Memory and Cognition*, *31*(1), 35-43.
- Mandler, G., Goodman, G. O., & Wilkes-Gibbs, D. L. (1982). The word-frequency paradox in recognition. *Memory and Cognition*, *10*(1), 33-42.
- Mangels, J. A., Picton, T. W., & Craik, F. I. (2001). Attention and successful episodic encoding: an event-related potential study. *Cognitive Brain Research*, *11*, 77-95.
- Mason, J. M. (1976). The roles of orthographic, phonological, and word frequency variables on word-nonword decisions. *American Educational Research Journal*, *13*(3), 199-206.
- McDaniel, M. A., & Geraci, L. (2006). Distinctiveness and memory. In R. R. Hunt & J. B. Worthen (Eds.), (p. 65-88). New York: Oxford University Press.
- McLaughlin, J. P. (1968). Recall and recognition measures of the von restorff effect in serial learning. *Journal of Experimental Psychology*, *78*(1), 99-102.
- Mecklinger, A. (in press). The control of long-term memory: Brain systems and cognitive processes. *Neuroscience and Biobehavioral Reviews*.
- Merritt, P. S., DeLosh, E. L., & McDaniel, M. A. (2006). Effects of word frequency on individual-item and serial order retention: Tests of the order-encoding view. *Memory and Cognition*, *34*(8), 1615-1627.

- Münter, T. F., Schiltz, K., & Kutas, M. (1998). When temporal terms belie conceptual order. *Nature*, 395, 71-73.
- Nelson, D. L., & McEvoy, C. L. (2000). What is this thing called frequency? *Memory and Cognition*, 28(4), 509-522.
- Otten, L. J., & Donchin, E. (2000). Relationship between p300 amplitude and subsequent recall for distinctive events: Dependence on type of distinctiveness attribute. *Psychophysiology*, 37, 644-661.
- Otten, L. J., Sveen, J., & Quayle, A. H. (2007). Distinct patterns of neural activity during memory formation of nonwords versus words. *Journal of Cognitive Neuroscience*, 19(11), 1776-1789.
- Paller, K. A., Kutas, M., & Mayes, A. R. (1987). Neural correlates of encoding in an incidental learning paradigm. *Electroencephalography and clinical Neurophysiology*, 67, 360-371.
- Polich, J. (1990). Probability and inter-stimulus interval effects on the p300 from auditory stimuli. *International Journal of Psychophysiology*, 10, 163-170.
- Polich, J., & Donchin, E. (1988). P300 and the word frequency effect. *Electroencephalography and clinical Neurophysiology*, 70, 33-45.
- Polich, J., Howard, L., & Starr, A. (1983). P300 latency correlates with digit span. *Psychophysiology*, 20(6), 665-669.
- Ruchkin, D. S., Johnson, R., Canoune, H., & Ritter, W. (1991). Event-related potentials during arithmetic and mental rotation. *Electroencephalography and clinical Neurophysiology*, 79, 473-487.
- Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high- and low-frequency words. *Memory and Cognition*, 18(4), 367-379.
- Rugg, M. D., Cox, C. J. C., Doyle, M. C., & Wells, T. (1994). Event-related potentials and the recollection of low and high frequency words. *Neuropsychologia*, 33(4), 471-484.
- Rugg, M. D., & Doyle, M. C. (1992). Event-related potentials and recognition memory for low- and high-frequency words. *Journal of Cognitive Neuroscience*, 4(1), 69-79.
- Saint-Aubin, J., & LeBlanc, J. (2005). Word frequency effects in immediate serial recall of pure and mixed lists: Tests of the associative link hypothesis. *Canadian Journal of Experimental Psychology*, 59(4), 219-227.

- Schmidt, S. R. (1991). Can we have a distinctive theory of memory? *Memory and Cognition*, *19*(6), 523-542.
- Squires, K. C., Donchin, E., Herning, R. I., & McCarthy, G. (1977). On the influence of task relevance and stimulus probability on event-related potential components. *Electroencephalography and clinical Neurophysiology*, *42*, 1-14.
- Squires, K. C., Wickens, C., Squires, N. K., & Donchin, E. (1976). The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science*, *193*(4258), 1142-1146.
- Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and clinical Neurophysiology*, *38*, 387-401.
- Steywers, M., & Malmberg, K. (2003). The effect of normative context variability on recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(5), 760-766.
- Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, *150*(3700), 1187-1188.
- Toglia, M. P., & Battig, W. F. (1987). *Handbook of semantic word norms*. Hillsdale, NJ: Erlbaum.
- Tse, C., & Altarriba, J. (2007). Testing the associative-link hypothesis in immediate serial recall: Evidence from word frequency and word imageability effects. *Memory*, *15*(6), 675-690.
- Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory and Cognition*, *18*(4), 380-393.
- Von Restorff, H. (1933). Über die wirkung von bereichsbildungen im spurenfeld. *Psychologische Forschung*, *18*, 299-342.
- Voss, J. L., & Paller, K. A. (2009). Remembering and knowing: Electrophysiological distinctions at encoding but not retrieval. *NeuroImage*, *46*, 280-289.
- Watkins, M. J., LeCompte, D. C., & Kim, K. (2000). Role of study strategy in recall of mixed lists of common and rare words. *Journal of Experimental Psychology*, *26*(1), 239-245.
- Wiswede, D., Rüsseler, J., & Münte, T. F. (2007). Serial position effects in free memory recall – an erp study. *Biological Psychology*, *75*, 185-193.
- Worthen, J. B. (2006). Distinctiveness and memory. In R. R. Hunt & J. B. Worthen (Eds.), (p. 133-156). New York: Oxford University Press.