LUNDS UNIVERSITET

INSTITUTIONEN FÖR PSYKOLOGI

Word frequency effects and inter-item associations. A study of the N400

Kristoffer Åberg

Kandidatuppsats vt 2006

Handledare: Mikael Johansson

Word frequency effects and inter-item associations. A study of the N400

Kristoffer Åberg

Lund University

Abstract

High frequency words are easier to recall compared to low frequency words in lists of only one frequency-range. This effect disappears when words of different frequencies are mixed in a list. Word frequency effects have often been assumed to rely on associations between words. The relationship between word frequency and associations was investigated by controlling associations using Latent Semantic Analysis (LSA) in an experiment of free recall. Event-related potentials (ERPs) were recorded and analysis focused on the N400 component. Behavioral results support the assumption that high frequency words have more associations than low frequency words while electrophysiological data do not. There was still a frequency effect in the N400 for lists of associated words. Results are not conclusive since the attempt to replicate word frequency effects in literature was unsuccessful and replication of results is necessary.

Keywords: Word frequency; N400; Episodic memory; Free recall; LSA; Latent Semantic Analysis; Semantic association

The word frequency effect in free recall is the find that words occuring often in a corpus of spoken or written text are easier to recall in pure (homogenous) lists compared to less frequent words (Fernandez et al., 1998; Gregg, Montgomery, & Castano, 1980; Hulme, Browne, & Morin, 2003; Watkins, LeCompte, & Kim, 2003, but see Terushkin, 2005).

When words of different frequencies are mixed in a list the probability of recall is increased for low frequency (LF) words while it is reduced for high frequency (HF) words (DeLosh & McDaniel, 1996; Fernandez et al., 1998; Gregg et al., 1980; MacLeod & Kampe, 1996; Terushkin, 2005). One explanation for word frequency effects have been the use of different rehearsal strategies since it has been found that HF-words are repeated more often, to more recent positions and more distributed compared to low frequency words (Tan & Ward, 2000). When controlling these factors by forcing participants to use the same rehearsal strategy for LF- and HF-words there is still an advantage in recall for HF-words, suggesting something else is driving the word frequency effect (Ward, Woodward, Stevens, & Stinson, 2003). Other explanations involve theories such as there being a limited amount of resources available at encoding, or that only the features of the word to be encoded matters independently of the surrounding words (see Hulme et al., 2003 for a discussion). Theories based on resource sharing cannot explain why mixing nonwords with HF-words improves recall for nonwords while recall for HF-words are not affected (Hulme et al., 2003, experiment 3). It is also clear that theories only taking into account the features of an item cannot explain why mixing HF- and LF-words changes probability of recall for both types of words. One set of theories that can account for most results concerning word frequency effects in free recall suggests that HF-words have more inter-item associations compared to LF-words (Gregg et al., 1980; Hulme et al., 2003). Gregg et al. let participants study lists of words where the inter-stimulus interval (ISI) between presentation of each word was varied. It was found that recall for HF-words in mixed lists was reduced compared to LF-words in the same lists, when ISI was increased (Gregg et al., 1980). According to an influential theory, inter-item associations between words play an important role at encoding when words occupy the short-term storage (STS) together (Glanzer, Koppenaal, & Nelson, 1972). Glanzer et al. argues that inter-item associations between words in the long term storage (LTS) are strengthened when words occupy the STS together. Since HF-words occur more often and in more contexts (Dennis, 1995), they have a higher probability of sharing the STS with other words and they will therefore have stronger (and possibly more) associations to others words. Increasing the ISI means words would spend less time together in the STS thus reducing any influence of inter-item associations. Since it was found that recall decreased for HF-words only, Gregg et al. concluded HF-words had more inter-item associations compared to LF-words. This conclusion goes well in hand with the search for associative memory model (SAM; Raaijmakers & Shiffrin, 1980). In SAM the probability of sampling an item for recall is based on the assumption that there are multiple cues available for participants. The probability of recall for an item is then a function of the multiplicative strengths of the various cues. If an item j has been recalled, then probability of recall for an item i is given by

$$P_s(i|C,j) = \frac{S_{Ci}S_{ij}}{\sum_k S_{Ck}Skj}$$

where S_{ij} is the associative strength between items j and i in the LTS, S_{Ci} is the associative strength between context C and item i. According to SAM, probability of recall should be similar when associations are controlled to be similar for HF- and LF-words.

Other similarities between effects of association and effects of frequency are that associated words have a higher probability of recall compared to words with lower associations and the same is true for HF-words compared to LF-words in pure lists (Gregg et al., 1980; Saint-Aubin, Ouellete, & Poirier, 2005). Response bursting is the phenomenon that associated words tend to become clustered at recall (Wingfield, Lindfield, & Kahana, 1998), this effect remains even when words are presented randomly suggesting it might rely both on co-occurence of words in STS and associations between words in the long-term storage thus providing a possible link between its frequency and associations (Kahana & Wingfield, 2000). This effect has also been found for HF-words thus suggesting they have more associations compared to LF-words (Bousfield, Cohen, & Whitmarsh, 1958; Cofer, Bruce, & Reicher, 1966). A related effect also believed to rely on HF-words having more associations than LF-words is that HF-words tend to be recalled more often in a free association task, possible because they use their associations as cues to generate other HF-words (Groot, 1989).

Similarities between frequency effects and associations can also be found in the concept of distinctivness (Schmidt, 1991). LF-words are considered being secondary distinct while isolated words (words that do not fit into the context where they are presented) are considered being primary distinct. The difference between secondary distinct and primary distinct words according to Schmidt, is that secondary distinct words are distinct in the LTS while primary distinct words are distinct in the STS. Both types of distinctivness yields similar behavioral patterns (such as distinct words showing less false alarms rates and higher hit rates in recognition), which suggests a link between the frequency of a word and in this case, lack of associations.

Even though there exist little empirical evidence supporting the assumption that HF-words have more inter-item associations it has been adopted by many theories and models (Anderson & Bower, 1972; Gillund & Shiffrin, 1984; Glanzer et al., 1972). It has been shown post-hoc that by using Latent Semantic Analysis (LSA) it is possible to connect associatons between words and word frequency effects. For example has it been found that probability of recall for a word was correlated with its association to a previously recalled word (Howard & Kahana, 1999; Howard & Kahana, 2002). In the analyis by Howard and Kahana (2002) it was also found that increasing the difficulty of a distractor task between presentation of words reduced the influence of association as a predictor of subsequent recall. These results provide a link between associations and frequency effects since they are similar to the results found by Gregg et al. (1980). Increasing the ISI between words reduces the predictive power in free recall of both word frequency and associations between words. According to the theory suggested by Glanzer et al. (1972), increasing the ISI would reduce the time two items spend together in the STS thus reducing encoding of associations between them, possibly because existing associations are not “discovered” (Glanzer, 1969). The fact that both frequency and associations became less good at predicting recall performance suggests they were influenced by the same mechanism: reduced strengthening of pre-existing associations in LTS.

Furthermore, LSA has been used to calculate associations between words in experiments investigating word frequency effects and it has been found that associations between words were highest in pure lists of HF-words, lowest in pure lists of LF-words, and intermediate in mixed lists thus suggesting another link between associations and the frequency of words (Hulme et al., 2003; Åberg, 2005). Using LSA to control associations between words of different frequencies in an experiment could provide important insights in the relationship between word frequency effects and inter-item associations.

Electroencephalogram (EEG) makes it possible to study changes in brain activity in milliseconds (Gevins, Smith, McEvoy, Leong, & Le, 1999). In order to increase the signal-to-noise ratio of the brains neurological responses it is common to extract chunks of EEG timelocked to events of interest and create averages of these chunks. These time-locked averages are called event-related potentials (ERPs) which contains deflections, or "bumps" when stimuli are processed (Luck, 2005). These deflections (or components), are commnly classified according to polarity, latency, and/or occurence. One of these components is a negative deflection occuring around 400 ms post stimulus called the N400. It was discovered in a study where subjects read sentences with anomalous endings which evoked the N400 (Kutas & Hillyard, 1980; see Kutas & Federmeier, 2000, for a review).

Different theories connect the N400 to the processing of words. Contextual integration is based on the assumption that the N400 reflects how easily a word is integrated into a context (Kutas & Federmeier, 2000). A context can be made up by only one word, a list of words, or a sentence (Kutas & Hillyard, 1980; Kutas & Federmeier, 2000; Rugg, 1990; Weisbrod et al., 1999). If a target word is associated with the context the amplitude of the N400 will be small reflecting easy integration of the word into the context (Chwilla, Hagoort, & Brown, 1998). Another theory suggests that the N400 reflects spreading activation in a network where words are represented by nodes (Collins & Loftus, 1975). As words become activated they will activate words similar in meaning by links between their respective nodes. If the representation of a word has been activated then processing of that word will be facilitated next time it gets activated, yielding a smaller N400. Both contextual integration and spreading activation can be argued to rely on sharing features in a network where overlap between features will facilitate processing. LSA provides a measure of similarity that can be interpreted as overlap between features, where features are represented as dimensions in a multidimensional space. Thus the more similar words are, based on a measure calculated using LSA, the amplitude of the N400 will be smaller possibly reflecting spreading activation in a network, or facilitated contextual integration. A third theory relating the N400 to word processing suggests that reading words creates predictions for words to follow (Federmeier & Kutas, 1999b). It has been found that the amplitude of the N400 is sensitive to the associations between an expected exemplar and a target word which could be interpreted as a comparison of features between exemplars and target word (Federmeier & Kutas, 1999a). Although the line between contextual integration and predictability of words might be a bit fuzzy, the assumption of predictability is that reading words will activate features of a predicted word or exemplar, while in contextual integration only features of preceding words will be activated. Thus predictability creates a set of predicted exemplars to which a target word is compared while for contextual integration the target word is compared to the features of the context. All of these theories could be argued to rely on associations between words where associated words evoke smaller N400 due to facilitated processing because of spreading activation, feature-overlap between a context and a target word, or feature-overlap between predicted examplars and a target word.

These theories can also explain the reduced N400 amplitude found for HF-words compared to LF-words (Fernandez et al., 1998; Osterhout, Bersick, & McKinnon, 1997). Assuming HF-words have more associations compared to LF-words then frequency effects in the N400 may rely on the same mechanisms as reviewed above.

As a summary, psychophysiological and behavioral data suggests a link between the frequency of a word and its associations. HF-words and associated words are easier to recall and they evoke smaller N400 amplitudes. LF-words and unrelated words are harder to recall and they evoke larger N400 amplitudes. Contrasting these two factors in an experiment would provide evidence of a link between them. LSA links word frequency to associations and provides an opportunity to control associations between words of different frequencies.

An experiment was created to investigate these relationships. Associations were controlled using LSA in lists of words of different frequencies. Due to the multitude of factors involved in the experiment it was decided to restrict the analysis of electrophysiological data to recalled words only. The downside of this approach is that the number of trials for each conditions will roughly be cut by half (if assuming 50% of words are recalled), yielding more noisy data. Another approach would be to collapse data from both recalled and forgotten words but this could mean perhaps finding an effect not related to word processing, but to differences at encoding for recalled and forgotten words. Adding a factor taking into account both recalled and forgotten words would solve this problem, but since subsequent memory effects in ERPs are beyond the scope of this report it was not included.

The hypothesis is that if word frequency effects rely on associations between words then controlling these associations will reduce any effects of word frequency. More specifically, recall and the amplitude of the N400 will be at an equal level for HF-words and LF-words when associations are controlled. Furthermore there will be no influence of list composition since it has been argued that this effect also relies on associations between words (Gregg et al., 1980; Hulme et al., 2003). A third prediction is that when associations are not controlled the amplitude of the N400 in mixed lists will be at an intermediate level due to intermediate levels of associations, as calculated by LSA (Hulme et al., 2003).

Method

Participants

29 participants (13 male, 16 female) served as participants (mean age 23.5, SD 3.6 years) and were recruited at classes, or via posters placed in the Lund University psychology department. All participants were healthy right-handers, native speakers of Swedish and they provided written, informed consent to participate in the study. Participants received a movie ticket for taking part in the experiment.

Stimuli and Design

Latent Semantic Analysis (LSA) was used to control associations between words. LSA is a method for deriving knowledge and learning from a body of text without using any pre-existing knowledge (Landauer & Dumais, 1997). LSA creates a multidimensional "semantic" space where each word is represented by a vector. By calculating a measure of similarity based on the projection of these vectors it is possible to find clusters of, or isolated words by setting tresholds of similarity. LSA has previously been used in post-hoc calculations of associations between words (Howard & Kahana, 2002; Hulme et al., 2003; Åberg, 2005). In order to control for, and investigate the influence of associations between words three conditions were created: associated, isolated, and natural lists of words. In the associated condition all words were extracted from clusters where all words were associated to any other word in the cluster. In the isolated condition all words were extracted that were not associated to any other words in the multidimensional space, and in the natural

Table 1: The experimental conditions. Factor ASSOCIATION is a between-groups factor, while factors FREQUENCY and COMPOSITION are within-group factors. In total there are 3 (associated, isolated, natural) x 2 (high frequency, low frequency) x 2 (pure lists, mixed lists) = 8 behavioral conditions in the experiment.

Factor			
ASSOCIATION	Isolated	Natural	Associated
FREQUENCY	High frequency	Low frequency	
COMPOSITION	Pure	Mixed	

condition there was no controlling of associations. The natural condition was included as a control condition in order to replicate word frequency effects in literature and as a baseline for any other effects achieved.

All words were extracted from the Stockholm-Umeå-Corpus (SUC), which consists of a representative collection of texts with a total of 1 million words (Källgren, 1998). Before being processed by LSA the words were lemmatized and divided into two frequency ranges with HF-words occuring more than 50 times in the corpus and LF-words occuring less than 4 times (thus yielding a frequency of 50, and 4 per million words). All of the words used in the experiment were nouns in singular form. Factors such as word length and concreteness were controlled and balanced between the conditions, words evoking a strong emotional response and **very** unusual words were removed. There was no manipulation of words chosen by LSA to be part of a cluster, even though intuitively some words should not belong to a cluster they were put in. The different experimental conditions are summed up in Table 1. In total there were 3 (associated, isolated, natural) x 2 (low frequency, high frequency) x 2 (mixed list, pure list) = 12 conditions which due to the nature of the experiment would be too long and demanding for participants to endure. It was therefore decided to make levels of association a between-group factor. In total there were 42 lists created for each level of association (20 pure lists, 20 mixed lists, and 2 practice lists). Every list consisted of 6 words and no word was repeated for any participant during the experiment. In the mixed lists, half of the lists started with a HF-word and the other half with a LF-word. Every second word in mixed lists was a HF-word.

Procedure

Participants were seated in a comfortable chair around 80 cm from the screen. They were informed about the experimental procedure and were instructed to focus on the present word and not use any memory strategies such as rehearsing preceding words or imagine items in a location. Subjects were also instructed not to move and to avoid blinking when words were presented. After instructions participants signed a written form of consent and the EEG equipment was applied.

At the start of the experiment and between each presentation of a word, a fixation cross appeared on the screen. Participants were told to focus on this. The words in the lists were presented sequentially for 1250 ms each with an inter-stimulus interval of 1000-1500 ms (including a time-jitter of 500-1000 ms). The words were presented in white on a black background. In order to reduce any effects of recency a distractor task was performed after each presentation of a list. After the last word of the list two numbers appeared on the screen, one between 50 and 100 and the other between 3 and 5. Subjects were instructed to count down in steps of the small number from the larger number. They were also instructed to do this out loud. Following the distractor task was a beep and a blank screen indicating the beginning of the recall phase. In the recall phase subjects were instructed to recall as many words as possible from the list and type the words down on a keyboard. The distractor task lasted for 10 seconds and participants had 20 seconds to type down recalled words. After this phase participants were shown a screen with instructions to press the space button to continue. During this period participants could take a pause, move, or blink if necessary.

EEG recording and analysis

During presentation of words, the electroencephalogram (EEG) was recorded from 38 electrodes placed according to the International 10/20 system (Jasper, 1958). Four additional electrodes (Silver/Silver-Chloride) were placed with adhesive tape in order to register eye movements. Two were placed at the outer canthi of the eyes and two were placed above and below the left eye. The tin electrodes were mounted in an elastic cap and activity was measured relative to reference electrodes placed behind the ears. One of the electrodes on the cap was connected to the ground. Impedances were held below 5k Ω. Biosignals were amplified and digitized with a sampling rate of 500 Hz using Neuroscan amplifiers and Acquire recording software (Neuroscan Inc. Sterling, VA).

After data had been recorded it was filtered using 0.1 Hz FIR high-pass filter and epoched -500 ms to 1500 ms relative onset of stimuli. Baseline correction was based on the 500 ms pre-stimulus activity. Prior to analysis all trials containing detectable, non-stereotypical artifacts were removed by visual inspection. Epochs containing stereotypical artifacts such as eye-blinks and eye-movements were kept in data because it is possible to correct for these using independent component analysis (ICA). After removal of bad epochs, EEG data was decomposed by ICA using the open source toolbox for matlab, EEGLAB (Delorme & Makeig, 2004). ICA is capable of isolating stereotypical artifacts such as eye-blinks, electrode drifts, and muscle activity and EEG data was corrected by rejecting ICA components reflecting such artifacts (Delorme & Makeig, 2004). The location of the N400 is commonly found in literature to be between 350-450 ms post stimulus and this was also approved by manual inspection. The maximal deflection at this time window was found at electrode CZ. CZ and surrounding electrodes (C3, C4, CPZ, and FCZ) were used in subsequent analysis of the N400. A mixed-factors repeated measures ANOVA was used in the preliminary analysis of ERP data for recalled words only, with between-groups factor ASSOCIATION (associated, isolated, and natural), and within-subjects factors FREQUENCY (high, low), COMPOSITION (mixed, pure).

Results

Behavioral data

Behavioral data as represented by probability of recall collapsed over all serial positions are presented in Figure 1 and Table 2. The reason for not including serial position as a factor in the analysis is that ERP analysis requires 20-30 trials for each factor investigated in order to reach a satisfactory signal-to-noise ratio. There are only 60 words for each condition, if probability of recall would be 50% then there would be only 30 (words) / 6 (serial positions) = 5 trials available for each condition. All results were corrected using Greenhouse-Geiser when the assumption of sphericity was violated. A mixed factor re-

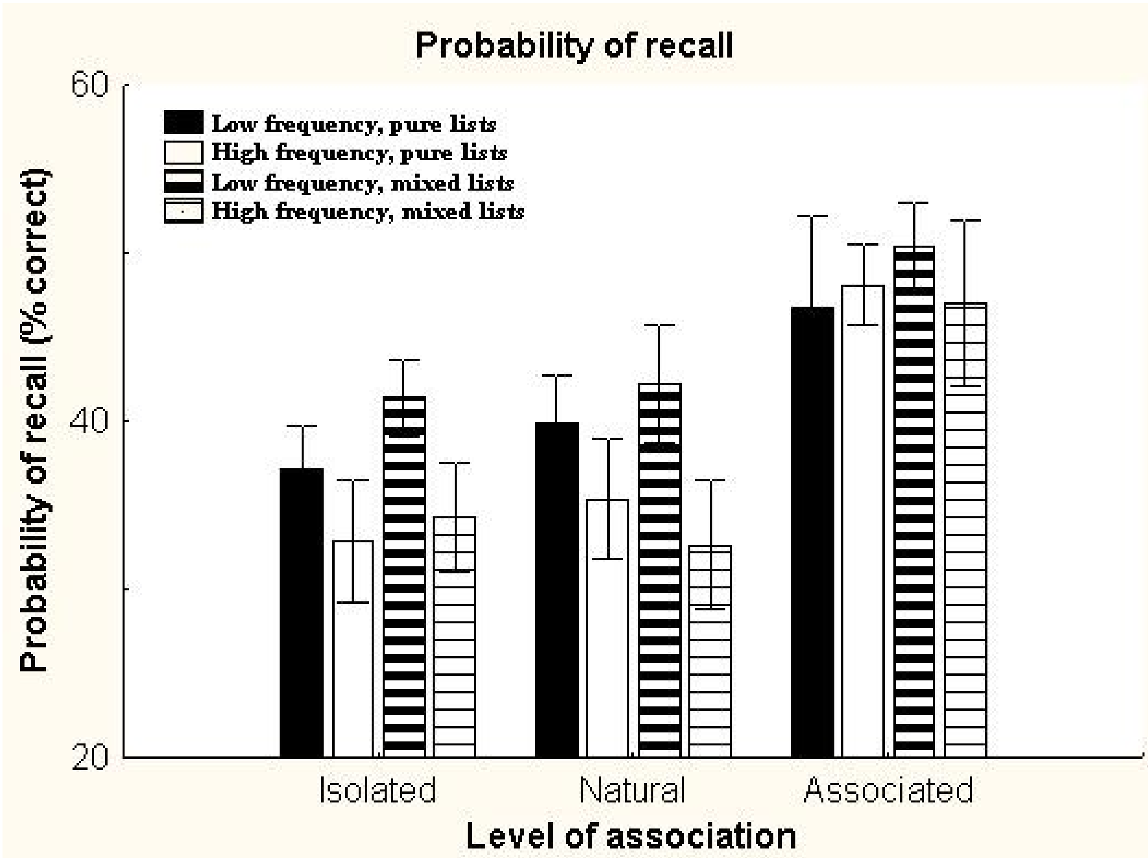

Figure 1. Probability of recall collapsed over all serial positions for each condition.

Table 2: Average probability of recall and Standard Error of the Mean (within parenthesis) collapsed over all serial positions for each condition.

	Pure lists		Mixed lists	
	LF-words	HF-words	LF-words	HF-words
isolated	.37 (.03)	.33 (.04)	.42 (.02)	.44 (.03)
natural	.40 (.03)	.35 (.04)	.42 (.04)	.33 (.04)
associated	.47 (.05)	.48 (.02)	.50 (.03)	.47 (.05)

peated measures ANOVA with between-group factor ASSOCIATION (natural, associated, isolated), and within-subject factors FREQUENCY (high, low), COMPOSITION (mixed, pure) revealed a main effect of FREQUENCY ($F(1,27) = 15.11$, $p < 0.002$) due to LF-words being recalled more often then HF-words. There was also a main effect of ASSOCIATION ($F(2,27) = 4.19$, $p < 0.03$) where associated words were recalled more often than both isolated and natural words. No other effects or interactions reached significance. Even though there was no interaction FREQUENCY*ASSOCIATION separate ANOVAs with the same within-subject factors as above was conducted in order to investigate the predictions that there should be no frequency effects when associations were controlled. For lists of associated words there was no frequency effect ($F(1,8) = 0.63$, $p > 0.44$), nor for isolated words ($F(1,9) = 4.05$, $p > 0.075$). In lists of natural words there was still an effect of frequency ($F(1,10) = 19.73$, $p < 0.002$), but there was no significant interaction COMPOSITION*FREQUENCY which indicates a failure to replicate word frequency effects found in literature. Thus, predicted results agree with experimental data when considering the reduction of frequency effects when associations were controlled. The failure to replicate word frequency effects might reflect confounds of concreteness effects, or a possible bias in the SUC-corpus. These assumptions will be elaborated upon in the discussion section.

ERP data

Grand averages for selected conditions averaged over electrodes FCZ, FZ, C3, C4, and CPZ can be seen in Figure 2. The grand averages for selected electrodes and conditions can be seen in Figures 3–5. As with behavioral data, all results were corrected using Greenhouse-Geiser when the assumption of sphericity was violated. The mixed factors repeated measures ANOVA with between-groups factor ASSOCIATION (associated, isolated, and natural), and within-subjects factors FREQUENCY (high, low), COMPOSITION (mixed, pure) for recalled words showed no main effect of factor COMPOSITION nor any interactions with this variable indicating unlike predicted that there was no modulation of the N400 in mixed lists for the natural condition. Due to factor COMPOSITION not explaining any variance in data, mixed and pure lists were collapsed across COMPOSITION in the subsequent analysis. ERP-data as measured by mean activity collapsed

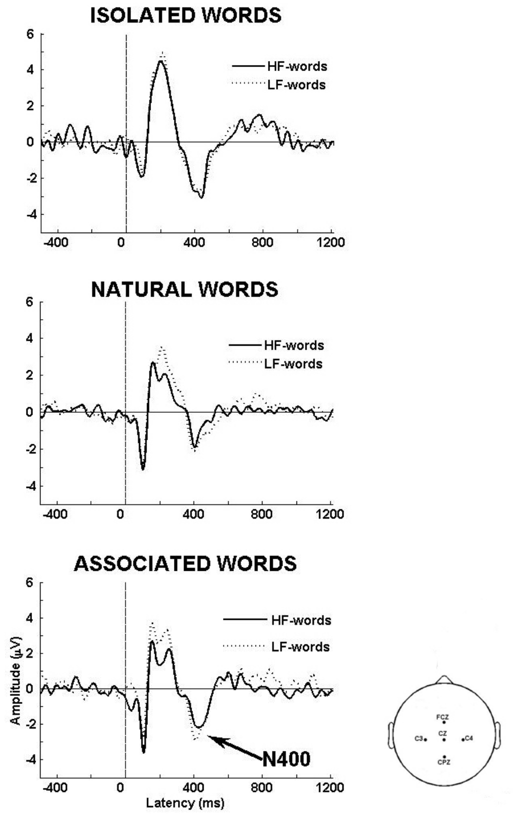

Figure 2. Grand averages of activity for recalled words averaged over electrodes FCZ, CZ, C3, C4, and CPZ.

over all serial positions and factor COMPOSITION in the time window 350-450 ms post stimulus and averaged over electrodes CZ, CPZ, FCZ, C3, and C4 are presented in Table 3. There was a main effect of FREQUENCY ($F(1,27) = 6.49$, $p < 0.02$), which was due to LF-words having more negative activation compared to HF-words. ASSOCIATION showed a trend ($F(2,27) = 2.54$, $p < 0.01$), where isolated words had more negative activation than both associated and natural words. Unlike predicted there was no interaction FREQUENCY*ASSOCIATION ($F(2,27) = 2.22$, $p > 0.13$), but in order to investigate the prediction of no frequency effects when associations were controlled separate ANOVAs were performed for each level of association with within-subject factors FREQUENCY (high, low). Natural words showed a frequency effect ($F(1,10) = 7.41$, $p < 0.03$), but in lists of isolated words there was no such effect ($F(1,9) = 0.03$, $p > 0.85$. Unlike predicted there was still a frequency effect in lists of associated words ($F(1,8) = 10.20$, $p < 0.02$). These

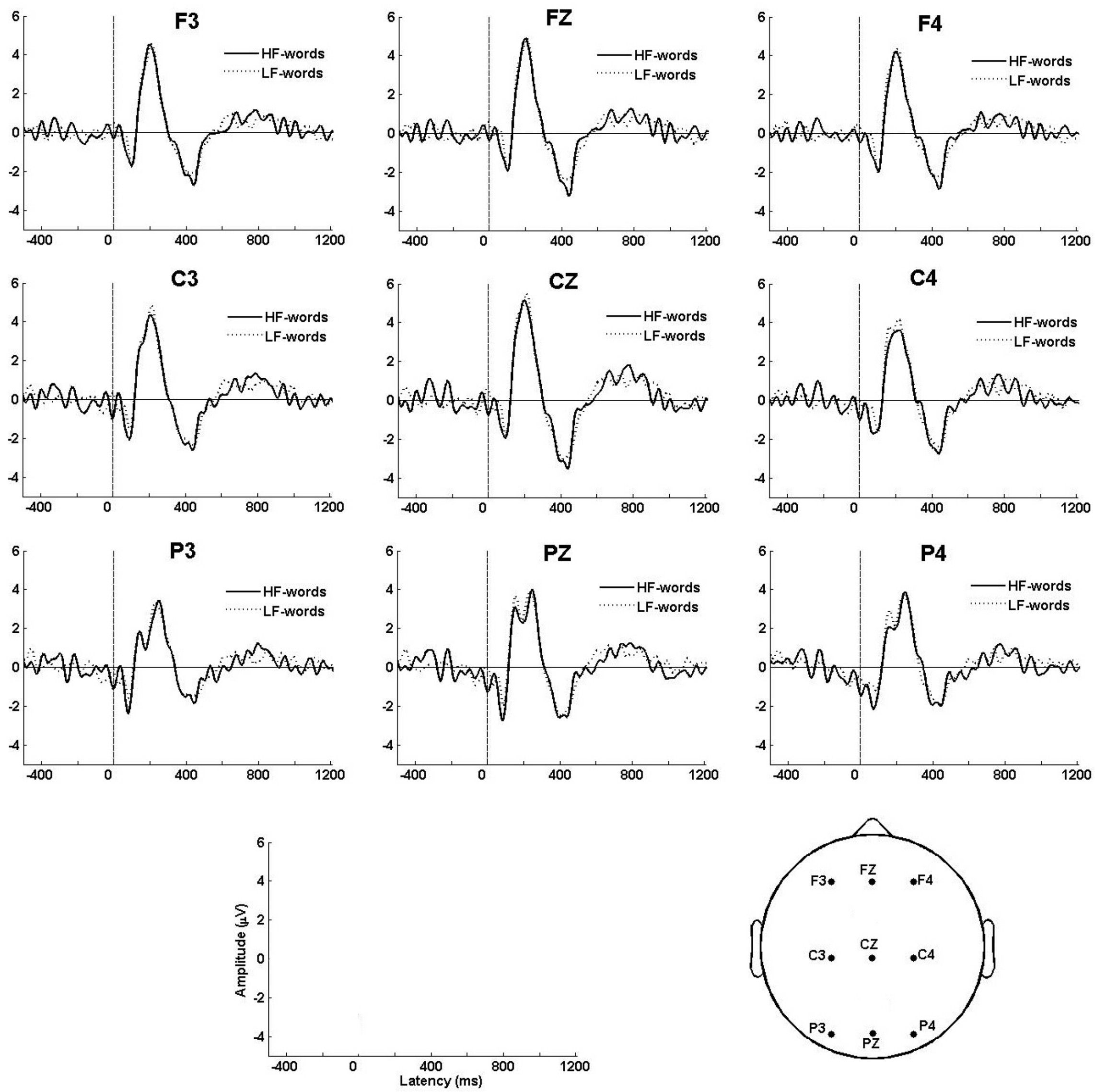

Figure 3. Isolated, recalled words. Grand averages for selected electrodes.

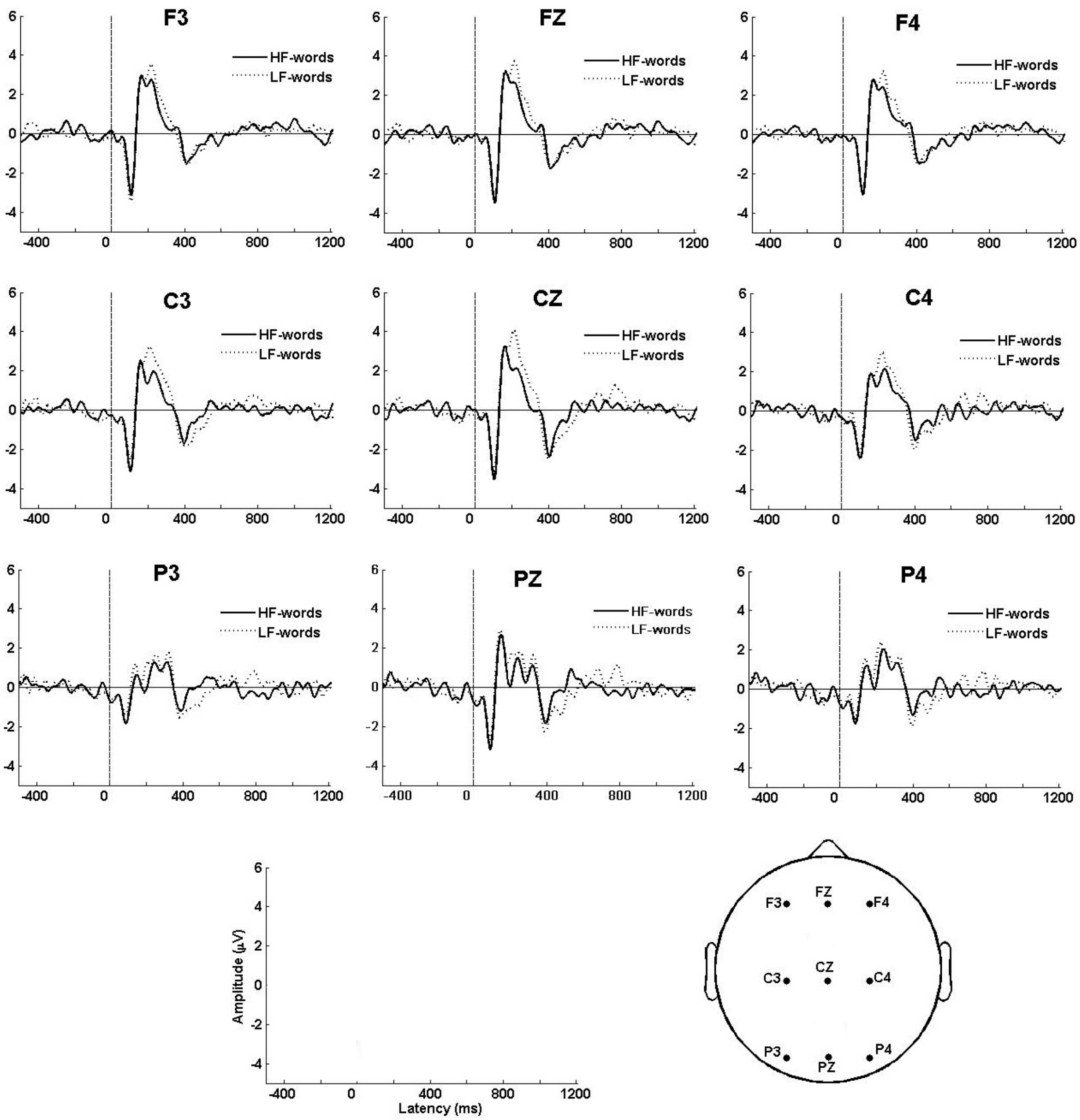

Figure 4. Natural, recalled words. Grand averages for selected electrodes.

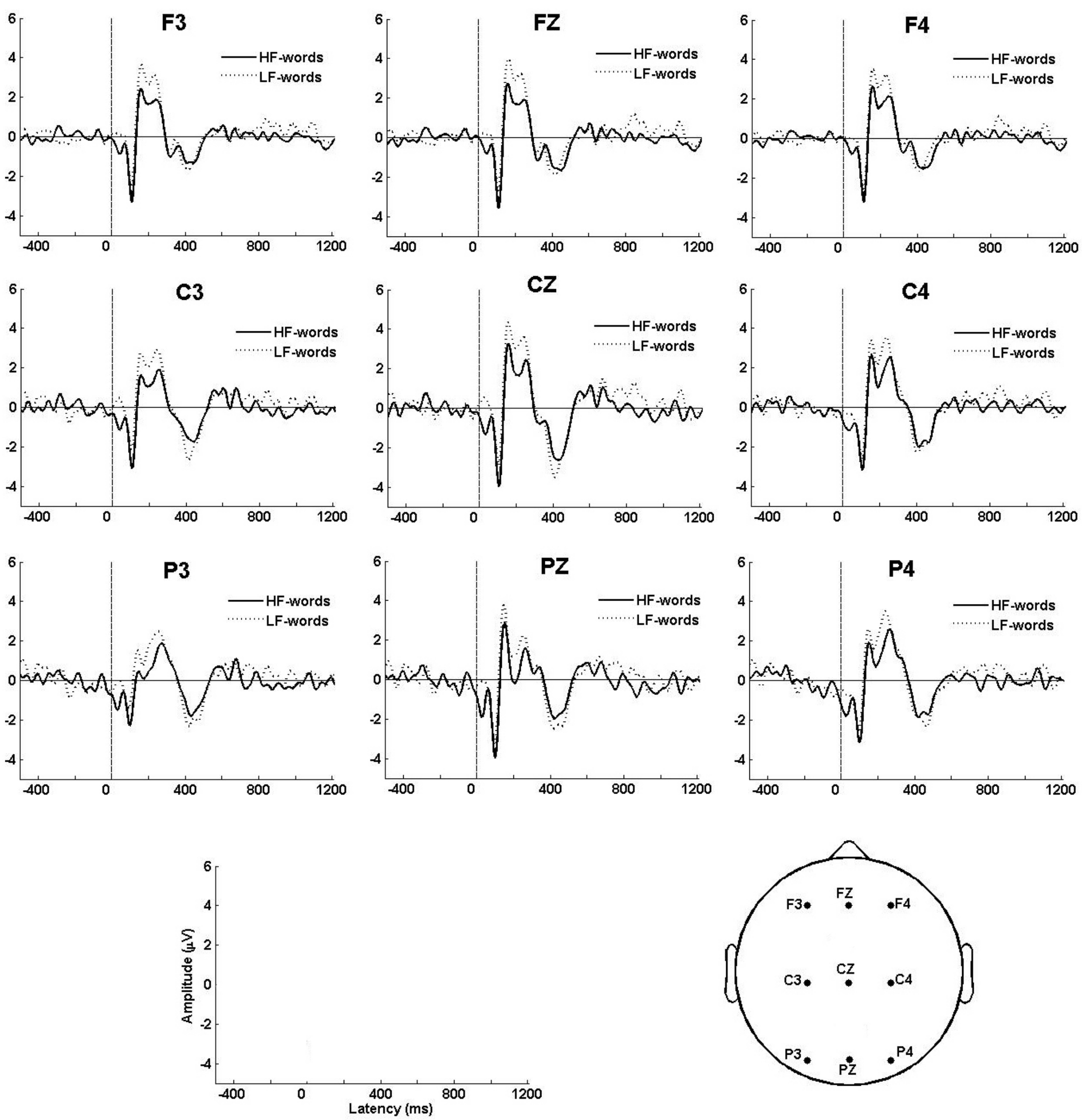

Figure 5. Associated, recalled words. Grand averages for selected electrodes.

Table 3: Mean amplitudes and Standard Error of the Mean (within parenthesis) in the timewindow 350-450 ms post stimulus averaged over electrodes FCZ, CZ, CPZ, C3, and C4.

	LF-words	HF-words
isolated	-2.15 (.29)	-2.20 (.41)
natural	-1.39 (.20)	-.98 (.17)
associated	-1.81 (.48)	-1.33 (.42)

results suggest that frequency effects might not be a result of associations, at least not as reflected in the amplitude of the N400.

Discussion

Frequency effects were present in natural lists of words in both behavioral and electrophysiological data. Controlling associations reduced frequency effects in behavioral data but only the amplitude of the N400 for lists of isolated words showed a reduction of frequency effects. This suggests creating associated lists of words did not modulate word frequency effects as reflected in the N400. Furthermore, interpreting data should be done with caution because the attempt to replicate word frequency effects in literature failed.

An attempt to explain word frequency effects in free recall was performed by using LSA to control for associations between words. Recall increased as associations were increased and differences in recall due to different frequencies were reduced in lists were associations were controlled (eg. associated and isolated lists of words). Word frequency effects in free recall has been assumed to rely on associations between words where HF-words have more associations than LF-words due to occuring more often and in more contexts (Gregg et al., 1980). This assumption have little support in empirical data of free recall, probably due to difficulties in controlling both word frequency and associations between words. LSA has previously been used post-hoc to explain word frequency effects (Hulme et al., 2003), but has never been used to directly control factors in an experiment. It might be argued that the reduction of word frequency effects in free recall was not due to controlling associations between words but due to the collection of words used in each condition. The first argument against this assumption is that all available words in the corpus were used as input to LSA and that no human intervention was used in order to “guide” the output of LSA into more appropriate lists of words. Even though some words chosen by LSA to be part of a list subjectively felt “out of context” they were kept in the list despite possible confounds of a von Restorff effect which is that words occuring out of context are easier to recall compared to words within the context (Fabiani & Donchin, 1995). There was also an equal opportunity for every word to appear in each condition (even though no isolated word could appear in the associated condition by definition). Secondly, using LSA to control associations resulted in an increased probability of recall for words in associated lists compared to isolated lists of words which replicates previous results (Saint-Aubin et al., 2005). A third argument is that LSA has previously been successfully used post hoc to explain effects in free recall involving associations (Howard & Kahana, 2002; Hulme et al., 2003; Åberg, 2005). Thus it can be argued that controlling associations using LSA successfully reduced word frequency effects in free recall. Another way to investigate if control of associations was successful in reducing frequency effects is analysis of ERPs in the time window of the N400.

The ERP-component N400 has been found to be sensitive to both associations and word frequency, where associated words and HF-words evoke a smaller N400 amplitude while the opposite is true for isolated words and LF-words (Fernandez et al., 1998; Kutas & Federmeier, 2000). Associated words and HF-words in the present experiment evoked lower amplitudes compared to LF-words and isolated words, confirming previous results. It was also predicted that if word frequency effects relied on associations between words then controlling the associations would reduce any differences between frequencies reflected in the amplitude of the N400. Results show that there were no frequency effects for isolated words while there was such an effect in the natural condition and for associated words which suggests that controlling associations using LSA was only partly successful.

This is a bit surprising because it has been shown that frequency effects are reduced when context is built up as when reading a sentence such that frequency effects are only present in words occuring at the beginning of a sentence and not at the end (Dambacher, Kliegl, Hofmann, & Jacobs, 2006; Kutas & Federmeier, 2000). Even though the present experiment did not involve sentences it could be argued that a context was built up when reading lists of associated words which should have reduced frequency effects. It might be that frequency effects are reduced for words at the end of the lists (serial position 5-6), but not for words occuring at earlier serial positions. As was argued before this could not be investigated in the present experiment due to the small amount of trials for each condition.

Theories of contextual integration, predictability, or spreading activation can all be argued to rely on overlap of features between different representations. A cat and a dog are more similar compared to a car and a dog because cats and dogs share more features (such as having a tail and fur, or being pets and animals). What LSA arguably does is creating a multidimensional space where each dimension can be considered an abstract feature of a word. Even though the dimensions in the LSA space cannot be considered concrete features, LSA connect theories of contextual integration and word frequency when creating its “feature-space”. It has been found that HF-words have a higher average value of similarity to all other words compared to LF-words (Åberg, 2005). This means that LSA provides a link between a words frequency and its associations to other words. Controlling associations after the space has been created would mean artifacts involving word frequency would also be takeninto account. It therefore seems plausible to believe that using LSA it would be possible to control word frequency effects in both behavioral performance and electrophysiological responses as measured by the N400. Unfortunately results show that this is only true for isolated lists of words since there is still an effect of frequency present for associated lists of words. One reason for this might be that associations due to frequency are somewhat different from more “semantic associations” as controlled by LSA.

Comparing studies of repetition priming and semantic priming could be argued to further dissociate effects of frequency and associations with the assumption that the N400 reflects associations between words (or overlap of features in an associative network). Repetition of a word and presenting associated words evoke similar behavioral and electrophysiological responses, such as lowered reaction times and a smaller N400 (Hill, Ott, & Weisbrod, 2005; Rugg, 1990). One theory trying to explain effects of priming is spreading activation which posits that words are represented as nodes in an associative network (Collins & Loftus, 1975). There exist evidence of a semantic (or associative) network but it is unclear whether such a network consist of nodes representing words, semantic features, or both (Deacon, Hewitt, & Tamny, 1998; Federmeier & Kutas, 1999b; Grose-Fifer & Deacon, 2004). By this account repeating a word would activate all features of the word twice while presenting associated words would activate some features of the second word twice. Note that this assumption might be a bit naïve since there are more processes involved such as processing a words physical form, linking it to meaning et cetera (Dambacher et al., 2006; Pulvermuller, 1999). Repeating a word would therefore activate the exact same processes twice while associated words might activate different processes prior to reaching an associative network where priming might be explained by spreading activation. The N400 is often believed to reflect post-lexical processing which might indicate activation of a semantic network as has been mentioned above Federmeier & Kutas, 1999a; Kutas & Federmeier, 2000). According to LSA, repeating an item and presenting two associated items would both induce priming irrespective of the frequency of the primed word. This is true for repetition priming (Rugg, 1990), where word frequency effects were reduced when repeting an item, but it is not true in the present study because word frequency effects reflected in the N400 are not reduced in lists of associated words. It therefore seems that effects of word frequency reflected in the N400 cannot be explained by HF-words having more associations than LF-words. Note that since there were no classical word frequency effects present it is hard to draw any conclusions of what was really manipulated by controlling associations using LSA. It should therefore not be concluded that HF-words does not have more associations compared to LF-words.

Behavioral data suggests that frequency effects were reduced using LSA to control associations between words. One influential model of episodic memory allows for different interpretations of the role of associations in free recall (SAM; Gillund & Shiffrin, 1984; Raaijmakers & Shiffrin, 1980). A basic assumption of SAM in free recall is that recalling a word (or image in SAM), correlates with its probability of being sampled based on available cues connected to associations in different ways. There are three different cues argued to influence word frequency effects in free recall: (1) self-sampling, (2) residual associations, and (3) relational encoding (Gillund & Shiffrin, 1984). The probability of sampling an item

i following a recently sampled item j can then be described as:

$$P_s(i|C,j) = \frac{S_{Ci}S_{ij}^{(1)}S_{ij}^{(2)}S_{ij}^{(3)}}{\sum_k S_{Ck}S_{kj}^{(1)}S_{kj}^{(2)}S_{kj}^{(3)}}$$

It has been suggested that word frequency effect depends on self-sampling and residual associations (Gillund & Shiffrin, 1984). Self-sampling occurs when words lack associations to other words in the studied list which means that the probability for a word to sample itself is increased which leads to an overall decrease in probability of recall. In pure lists of LF-words, probability of recall would be decreased because LF-words have less associations to other words and thus would sample themselves more often compared to HF-words which would have a smaller probability of sampling themselves in pure lists of HF-words. According to this assumption one explanation why controlling associations reduces frequency effects is that when associations are controlled the probability of self-sampling is equal for LF- and HF-words which means there is no longer an advantage for HF-words in pure lists. Residual associations is another mechanism used to explain word frequency effects where it is assumed that HF-words have stronger associations to any words compared to LF-words which is a result of encoding processes accumulated over a long time. Thus if a HF-word has been sampled (recalled), the probability of recovering any other word is increased compared to if a LF-word has been sampled. This explains why HF-words have a higher probability of recall in pure lists. Even though these mechanisms are often treated as independent they could be argued to reflect a common basic mechanism: associations between items. If these two mechanisms were totally independent they would not be able to explain why controlling associations in lists of isolated words reduces frequency effects since there should still be an advantage for HF-words due to residual associations which are not influenced by experimental manipulations to the same extent as self-sampling. One possible explanation for this connects LSA, word frequency effects and associations in a nice way.

A third mechanism of SAM is that of relational encoding which suggests that as participants encode words they will connect features of words and use these features as cues in order to sample these words at later recall. It was suggested by Glanzer et al. (1972) that as words occupy the STS together their associations in LTS are strenghtened. Assuming that relational encoding does not require any conscious “discovery” of associations or feature overlap between words it could be argued that relational encoding is what drives the word frequency effects since LSA provides a link between word frequency and associations between words. LSA also connects relational encoding to residual associations since residual associations are related to the frequency of a word and relational encoding is dependent on associations, or feature overlap between words. Controlling associations using LSA would control residual associations, which according to LSA would just be an artifact due to the frequency of a word. Residual associations would then only matter in lists of natural words and this would also explain why increasing the ISI between words reduced the predicitive power of both word frequency and associations between words (Gregg et al., 1980; Howard & Kahana, 2002).

In lists of natural words, there was no advantage for HF-words in pure lists, nor was there an interaction between frequency and list composition which is a common find. This indicates controlling factors such as concreteness failed. There was an attempt to control for concreteness in the reported experiment but due to the limited size of the corpus it was not possible to achieve a perfect balance between conditions when also controlling for factors such as word frequency, word length, and associations between words. The only factor not perfectly balanced between conditions was concreteness where LF-words were rated more concrete than HF-words. Concrete words have been reported to be easier to recall compared to abstract words (Nittono, Suehiro, & Hori, 2002), which might explain why there is no advantage for HF-words in pure lists (but it has also been reported that effects of concreteness only influenced recall of HF-words Bosco, 2005).

Besides consisting of fewer words compared to other corpuses commonly used to study word frequency effects, the SUC-corpus might also suffer from a bias not found in other corpuses. A reason for believing this is that another study also report an abscence of word frequency effects using the SUC-corpus (Sikström, Kallioinen, Gredebäck, & Rosander, Submitted). It has been reported that LF-words tend to occur frequently in a few contexts while HF-words occur seldom in many contexts (Dennis, 1995). The SUC corpus consist of less than 1000 texts ranging from book chapters to food-receipts. This composition of texts, together with the process of lemmatizing words might create HF-words from words that might be of low, or middle frequencies in other corpuses thus influencing recall performance.

One possible confound of data might be effects of recency, which is the find that words at the end of lists are recalled more often than words in the middle of lists (Greene, 1986). This effect is believed to rely on different mechanisms compared to encoding of words earlier in the list which would confound both behavioral and electrophysiological data. This would be a problem for the present study if the N400 differs from words recalled by recency and words recalled due to encoding. If one condition contains more trials from words being recalled due to recency, then the distributions would be skewed possibly resulting in more variance which can be explained by a factor based on serial position. One way to answer these questions would be to perform another experiment with more participants per condition and take into account when in a list a word is presented. As was mentioned no such analysis of ERP-data is possible in the present study due to very few trials at each serial position.

Due to these possible confounds results are not conclusive in favor of word frequency effects in free recall being explained by associations between words and replication of results is necessary when using a more strict control of concreteness and possibly another corpus.

In conclusion, controlling associations between words reduced word frequency effects in free recall. The reason for this is believed to be because controlling associations using LSA also controls associations driving the word frequency effect. By controlling these associations there are no longer any advantage for HF-words at encoding and probability of recall will be at an equal level for HF-words and LF-words. Electrophysiological data does not support the assumption that HF-words have more associations compared to LF-words since there was still a difference in amplitude of the N400 between HF- and LF-words in associated lists of words. The classical word frequency effect was not present in the experiment which makes data difficult to interpret and there is a need to replicate the results due to possible confounds with factors such as concreteness, recency effects, and a possible bias in the corpus where words were derived from.

References

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review*, *79*, 97–123.

Åberg, K. (2005). *Explaining the word frequency effect by application of a computational model of semantic memory and a suggested improvement for an existing model of episodic memory.* Unpublished master's thesis, Department of Cognitive Science, Lund University.

Bosco, F. A. (2005). *Effects of test expectancy, word frequency, and word concreteness on encoding workload.* Unpublished master's thesis, Montana State University.

Bousfield, W. A., Cohen, B. H., & Whitmarsh, G. A. (1958). Associative clustering in the recall of words of different taxonomic frequencies of occurrence. *Psychological Reports*, *4*, 39–44.

Chwilla, D., Hagoort, P., & Brown, C. (1998). The mechanism underlying backward priming in a lexical decision task: Spreading activation versus semantic matching. *The Quarterly Journal of Experimental Psychology*, *51A*, 531–560.

Cofer, C. N., Bruce, D. R., & Reicher, G. M. (1966). Clustering in free recall as a function of certain methodological variations. *Journal of Experimental Psychology*, *71*, 858–866.

Collins, A., & Loftus, E. (1975). A spreading activation theory of semantic processing. *Psychological Review*, *82*, 407–428.

Dambacher, M., Kliegl, R., Hofmann, M., & Jacobs, A. M. (2006). Frequency and predictability effects on event-related potentials during reading. *Brain Research*, *1084*, 89–103.

Deacon, D., Hewitt, S., & Tamny, T. (1998). Event-related potential indices of semantic priming following an unrelated intervening item. *Cognitive Brain Research*, *6*, 219–225.

Delorme, A., & Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *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*, 1135–1146.

Dennis, S. (1995). The impact of the environment on the word frequency and null list strength effects in recognition memory. *Noetica*, *1*, 1.

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*, 224–240.

Federmeier, K. D., & Kutas, M. (1999a). Right words and left words: electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research*, *8*, 373–392.

Federmeier, K. D., & Kutas, M. (1999b). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, *41*, 469–495.

Fernandez, G., Weyerts, H., Tendolkar, I., Smid, H., Scholz, M., & Heinze, H.-J. (1998). Event-related potentials of verbal encoding into episodic memory: Dissociation between the effects of subsequent memory performance and distinctiveness. *Psychophysiology*, *35*, 709–720.

Gevins, A., Smith, M. E., McEvoy, L. K., Leong, H., & Le, J. (1999). Electroencephalogram imaging of higher brain function. *Philosophical Transactions: Biological Sciences*, *354*, 1125–1134.

Gillund, G., & Shiffrin, R. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *19*, 1–65.

Glanzer, M. (1969). Distance between related words if free recall : Trace of the sts. *Journal of Verbal Learning and Verbal Behavior*, *8*, 105–111.

Glanzer, M., Koppenaal, L., & Nelson, R. (1972). Effects of relations between words on short-term storage and long-term storage. *Journal of Verbal Learning and Verbal Behavior*, *11*, 403–416.

Greene, R. L. (1986). Sources of recency effects in free recall. *Psychological Bulletin*, *99*, 221–228.

Gregg, V. H., Montgomery, D. C., & Castano, D. (1980). Recall of common and uncommon words from pure and mixed lists. *Journal of Verbal Learning and Verbal Behavior*, *19*, 240–245.

Groot, d. A. M. B. (1989). Representational aspect of word imageability and word frequency effects as accessed through word association. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 824–845.

Grose-Fifer, J., & Deacon, D. (2004). Priming by natural category membership in the left and right cerebral hemispheres. *Neuropsychologica*, *42*, 1948–1960.

Hill, H., Ott, F., & Weisbrod, M. (2005). Soa-dependent n400 and p300 semantic priming effects using pseudoword primes and a delayed lexical decision. *International Journal of Psychophysiology*, *56*, 209–221.

Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 923–941.

Howard, M. W., & Kahana, M. J. (2002). When does semantic similarity help episodic retrieval? *Journal of Memory and Language*, *46*, 85–98.

Hulme, C., Browne, G., & 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.

Jasper, H. H. (1958). The ten-twenty electrode system of the international federation. *Electroencephalogram in Clinical Neurophysiology*, *10*, 371–375.

Kahana, M. J., & Wingfield, A. (2000). A functional relation between learning and organization in free recall. *Psychonomic Bulletin & Review*, *7*, 516–521.

Källgren, G. (1998). *Documentation of the stockholm-umeå corpus.* (Department of Linguistics, Stockholm University)

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*, 463–469.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203–205.

Landauer, T. K., & Dumais, S. T. (1997). A solution to plato's problem: The latent semantic analysis theory of aquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211–240.

Luck, S. J. (2005). *An introduction to the event-related potential technique.* The MIT Press: MA.

MacLeod, C. M., & Kampe, K. E. (1996). Word frequency effects on recall, recognition, and word fragment completion tests. *Journal of experimental psychology: Learning, Memory, and Cognition*, *22*, 132–142.

Nittono, H., Suehiro, M., & Hori, T. (2002). Word imageability and n400 in an incidental memory paradigm. *International Journal of Psychophysiology*, *44*, 219–229.

Osterhout, L., Bersick, M., & McKinnon, R. (1997). Brain potentials elicited by words: word length and frequency predict the latency of an early negativity. *Biological Psychology*, *46*, 143–168.

Pulvermuller, F. (1999). Words in the brain's language. *Behavioural & Brain Sciences*, *22*, 253–336.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1980). Sam: A theory of probabilistic search of associative memory. *The psychology of learning and motivation*, *14*, 207–262.

Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high- and low-frequency words. *Memory & Cognition*, *18*, 367–379.

Saint-Aubin, J., Ouellete, D., & Poirier, M. (2005). Semantic similarity and immediate serial recall: Is there an effect on all trials? *Psychonomic Bulletin & Review*, *12*, 171–177.

Schmidt, S. R. (1991). Can we have a distinct theory of memory? *Memory & Cognition*, *19*, 523–542.

Sikström, S., Kallioinen, P., Gredebäck, G., & Rosander, K. (Submitted). *Erp and serial position curves during encoding of high and low frequency words.* (Submitted)

Tan, L., & Ward, G. (2000). A recency-based account of the primacy effect in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1589–1625.

Terushkin, V. (2005, May). *Human episodic memory from a dual behavioral/electrophysiological perspective.* Senior Honors Thesis.

Ward, G., Woodward, G., Stevens, A., & Stinson, C. (2003). Using overt rehearsals to explain word frequency effects in free recall. *Journal of Experimental Psychology*, *29*, 186–210.

Watkins, M. J., LeCompte, D. C., & Kim, K. (2003). Role of study strategy in recall of mixed lists of common and rare words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 239–245.

Weisbrod, M., Kiefer, M., Winkler, S., Maier, S., Hill, H., Roesch-Ely, D., et al. (1999). Electrophysiological correlates of direct versus indirect semantic priming in normal volunteers. *Cognitive Brain Reserach*, *8*, 289–298.

Wingfield, A., Lindfield, K. C., & Kahana, M. J. (1998). Adult age differences in the temporal characteristics of category free recall. *Psychology and Aging*, *13*, 256–266.