UC Berkeley
UC Berkeley PhonLab Annual Report

Title
Neural Basis of the Word Frequency Effect and its Relation to Lexical Processing

Permalink
https://escholarship.org/uc/item/4th427hg

Journal
UC Berkeley PhonLab Annual Report, 8(8)

ISSN
2768-5047

Author
Kang, Shinae

Publication Date
2012

DOI
10.5070/P74th427hg

eScholarship.org Powered by the California Digital Library
University of California

Neural basis of the Word Frequency Effect and Its Relation to Lexical Processing

Shinae Kang *

January 21, 2013

1 Introduction

A number of experimental results have shown that in terms of cognitive processing common words – words that occur frequently - differ from uncommon words. People perceive common words more accurately and more quickly when listening (Savin, 1963) and also name them more quickly (Oldfield and Wingfield, 1965) while speaking. Also, people produce the common words in more reduced and shortened forms than uncommon words (Whalen, 1991; Gahl, 2008). These results indicate a certain processing difference between the common (or frequent) words and the uncommon (or infrequent) words. This phenomenon has been extensively studied in the psycholinguistic literature, in particular for speech production, as one way of probing the cognitive architecture. Current models in single word production generally agree that the production process consists of multiple cognitive actions (Dell, 1986; Levelt et al. 1999). Broadly speaking, they include conceptulization, retrieval of syntactic and semantic information from the mental lexicon, retrieval of phonological form, assembly of sounds into syllables (syllabification), and finally implementation of speech motor plan in terms of commands to specific muscles to execute the articulation. It is possible that any one, or all, of these activities could be affected by word frequency. Indeed several studies have pointed to an effect of word frequency at different stages of speech production (see e.g. the contrasting accounts of Jescheniak & Levelt, 1994; Gahl, 2008, details on Section 2). However, studies seeking neural evidence of the processing difference by word frequency have so far yielded inconsistent results that cannot be uniformly accounted for. The present study aims to fill this gap by finding a more reliable neural basis of the effect of word frequency using high density intracranial recordings during word reading.

In addition, this study analyzes both one- and two-syllable words in order to observe a possible processing difference caused by the number of syllables. Since several studies suggest that word frequency affects processing of words differently depending on the number of syllables (Balota& Chumbley, 1985; Jescheniak& Levelt, 1994), it might be relevant to

---

[*] Data from this study was collected and made available by the lab of Dr. Edward Chang at UCSF. I appreciate suggestions and feedbacks from Edward Chang, Kristofer Bouchard, Nima Mesgarani, Angela Ren, Keith Johnson, Emily Cibelli, and Susanne Gahl, and members of the Chang Lab at UCSF and the Phonology Lab at UC Berkeley.

look at both one-syllable and two-syllable words together to understand how they differ from each other. Also, in order to capture the nature of word frequency more accurately, it is important to observe the neural structures of the different processing pattern–how frequency is physically realized at the neural level. However, studies on the neural basis of the word frequency has been inconclusive so far, partially because of insufficient data taken from either only spatial (Carreiras et al., 2006; Graves et al., 2007) or only temporal (Sahin et al., 2009; Laganaro et al., 2009; Strijkers et al. 2010) dimensions. The intracranial recording that the current study uses is direct cortical electrocorticography (ECoG), which captured the spatiotemporal dynamics of brain activities while a subject read one- and two-syllable words of different frequency. This study therefore provides a better picture of the underlying dynamics of the frequency effect, providing a spatial mapping of brain activity locked in a finer-grained temporal point. To preview the result briefly, word frequency effect was found for both one- and two- syllable words, but interestingly the spatio-temporal patterns differed according to the number of syllables. This suggeststhat there are different frequency representations for one-syllable words and two-syllable words.

The outline of the paper is follows: Section 2 elaborates on the the background of the study. Sections 3 and 4 detail the procedure for collecting and analyzing data. Section 5 presents some of the preliminary data analyzed so far, comparing the patterns shown in different regions of the brain. This is followed by a tentative discussion in Section 6. Section 7 discusses the drawbacks and future directions of this study.

2 Background

2.1 Speech production and locus of the word frequency effect

A large body of psycholinguistics literature has explored the locus of the word frequency effect in speech production; how and where word frequency plays a role within the functional architecture of word production. Despite some variations, current cognitive models of single-word production broadly assumes three main processes: 1) retrieval of conceptual information 2) formulation of words that carry the information and 3) articulation (Dell et al., 1986; Levelt et al. 1999).The second process, formulation of words, is further composed of selecting an entity that has syntactic and semantic information (lemma) and retrieving a phonological form corresponding to the selected lemma. Phonological forms include segments (e.g. phoneme) and metrical information (syllable, stress etc.) and the process of retriving them is referred as phonological encoding. Following phonological encoding comes the third process, phonetic encoding where the phonological form is uttered through the articulatory planning stage and the actual articulation.

It was found that word frequency can affect one or multiple stages in the production model. One group of researchers tested if articulatory planning and articulation stage are affected by word frequency (Balota & Chumbley, 1985; Jescheniak & Levelt, 1994). The studies concern whether the locus of word frequency is at the lexical or post-lexical processing level. Assuming that articulatory processes occur subsequently to meaning and form retrieval, articulation and its planning were considered to occur at the post-lexical level (Balota & Chumbley, 1985). These studies were interested in whether word frequency affects only earlier stages in production prior to articulatory planning or also influences the articulation stages as well.

Balota and Chumbley (1985) exploited a delayed naming task to test the effect of word frequency on articulatory processes. In their Experiment 3, the subjects first saw a word and after some delay pronounced the word in response to a 'go' signal. The authors' argued that if the delay is long enough, processes that are assumed to occur early (e.g. lemma selection and phonological form selection) should be completed by the time they articulate the word. Hence, any remaining difference in the naming latency would be attributable to the subsequent processes; phonetic encoding. They found that the difference in the naming latency made by word frequency decreased as the length of delay increased, but it did not diminish completely. With this result they concluded that part of the frequency effect on naming latency is localized in the articulation stage. However they remained agnostic about the impact of frequency on the subsequent processes prior to articulation.

Jeschniak and Levelt (1994) reported an opposing result. They conducted seven experiments, each of which tested if word frequency affects each process in the production model including articulation. In their Experiments 3 and 7 where experiment conditions were deliberately controlled to be the same as those in Balota and Chumbly (1985), they found no significant difference in naming latency for frequent and infrequent words when long delays (1000 msec, 1600 msec for experiment 3 and 7, respectively) were placed between the stimuli and the go signal. As a result, the phonetic encoding process was excluded from being a potential locus of word frequency. Building on other evidence found in their other experiments, they argued that the process where the phonological form is retrieved is the sole locus at which word frequency information is incorporated, contra the finding by Balota and Chumbley (1985). Jeschniak and Levelt (1994) tried to account the inconsist results by pointing to two related factors: 1) type of stimuli and 2) the duration of delays. Balota and Chumbley (1985) used multisyllabic stimuli and shorter delays whereas Jeschniak and Levelt (1994) used monosyllabic stimuli. Based on these methodological differences,, Jeschniak and Levelt (1994) posited that multisyllabic words would inherently require greater a degree of phonetic encoding that a short delay would not be sufficient for the encoding to complete entirely.

This account, however, fails to fully explain the cause of the inconsistency. One possible criticism comes from the assumptions concerning strict temporal sequence between stages. Without knowing exactly how the processes in the productions models come about, including whether these stages are to some extent parallel or in cascading relationship with one another (Sinai, 2008), the determination of long and short delays is arbitrary however accurate the studies attempted. Since the delayed naming task also assumes strict sequentiality of the subprocesses, it is not an accurate method to measure any fine-grained dynamics of different stages. Another criticism comes from the assumption that one-syllable words words and multisyllabic words undergo the same processes, but differ only in terms of magnitude of articulatory load. It could be that there is more than one uniform representation of word frequency depending on the number of syllables.

In addition to distinguishing between frequency effects during lexical and articulatory processing (as just discussed above), some researchers have delved specifically into two distinct steps *within* the processing at the lexical level: the retrieval of lemma vs the retrieval of phonological form (Kittredge et al., 2008). Jescheniak & Levelt (1994) claimed earlier that it was only the (phonological) form frequency that is central to the lexical frequency not lemma. However, this finding was challenged by evidence found with the durations of homophones (Gahl, 2008). Shorter word duration for frequent words was assumed to be an articulatory benefit that frequent words have over less frequent words (Balota & Chumbley, 1985). Gahl (2008) particulatly tested if this articulation benefit is solely due to articulatory routinization (increased motor fluency) or also modulated by higher-level lexical information such as lemma frequency. Gahl (2008) measured durations of words that have the same phonological form but different meaning. If word frequency only affected the phonological form where all phonological encoding had occurred already, infrequent words that have frequent form counterparts would be shortened as much as the the frequent words since they share the same phonological forms. The study found that when all other potential factors are controlled for, the frequent forms were shorter than their homophonic infrequent counterparts. Gahl (2008) concluded that lemma frequency is a relavant factor in word duration difference in homophones, supporting the view that word frequency is a property that can pervade to all levels of production.

As briefly shown, the literature have focused on various types of behavioral contrasts for frequent and infrequent words in order to answer different parts of the question regarding the representation of frequency information within the production model. So far the results seem to be inconsistent sometimes, if not comparable to one another at all. One of the potential reasons is the different frequency representations for words of different syllable length, as was suggested by contrasting results by Balota and Chumbley (1985) and Jescheniak and Levelt (1999). If one-syllable words and two-syllable words are processed differently because of the different degree of syllabification or related process, the seemingly-inconsistent result can be accounted in the consistent manner.

2.2 Neuroimaging studies on the word frequency effect

Studying the neural basis of word frequency is a way to better understand the cognitive processing at the implementation level. Recent development in brain imaging/electrophysiological techniques has made it possible to observe the neural signals in more detail by mapping the neural firing at the different brain regions as they change over time. Because functions of each brain regions are relatively independent for a different cognitive action, although still interconnected to each other to some extent, observing how activation is carried out in each region and how it differs for frequent and infrequent words will help clarify complex mechanisms underlying the effect of word frequency. According to the most comprehensive meta-analysis to date, different brain regions are indeed activated at different times over the course of production (Indefrey & Levelt, 2004; Indefrey, 2011). The typical pattern of activation roughly corresponds to specific cognitive actions postulated in the production models by Levelt et al. (1999).

The early neuroimaging studies collectively worked on mapping out each brain region and linking them to each cognitive function posed in the production model (Indefrey, 2011). The biggest finding from this mapping approach is that a process traditionally considered to be indivisible are in fact realized in multiple distinct brain regions. For example, researchers found that phonological encoding in the psycholinguistics production models is associated with two different regions, the right superior temporal gyrus (pSTG) possibly corresponding to phonological form retrieval, and the left inferior frontal gyrus (IFG) for syllabification. The other processes in the production model were also confirmed with different brain regions activated for each function (Indefrey & Levelt, 2004).

Beyond basic mapping, the focus of neuroimaging studies has expanded to hypothesis-testing, e.g., figuring out the neural correlates of different lexical variables and their effect on the processing stream in detail. Such studies aimed to provide empirical evidence on the effect of the various lexical peroperties e.g. neighborhood density (Prabhakaran et al. 2006), word frequency (Prabhakaran et al. 2006; Graves et al., 2007; Laganaro et al., 2009; Strijkers et al. 2010) and so forth. Each study used diffrent kinds of neuroimaging (e.g.fMRI) or neurophysiological techniques (e.g. EEG). These techniques provided different aspect of the research question. For example for the same topic of word frequency, a functional imaging study (Prabhakaran et al. 2006; Graves et al., 2007) offered a detailed account on spatial dimensions of brain activity, but the electrophysiological studies using EEG provided results in a temporal dimension like the time course of the activation (Laganaro et al., 2009; Strijkers et al. 2010) . However, the neural correlate of the word frequency could not be concluded easily due to substantial inconsistencies among the results caused by the inherent limitation of each type of technique being blind to the other dimensions.

Functional magnetic resonance imaging (fMRI) observes the activation level in different regions while subjects perform specific tasks by measuring blood oxygen level when they are responding to certain tasks. Prahabkaran et al. (2006) explored the neural correlates of the word frequency effect on word recognition with a lexical-decision task. From a pool of high and low frequency words, subjects determined if a stimulus was a word or not, during which the subjects' cortical activation was measured by event-related fMRI. Prahabkaran et al. (2006) found that the activation pattern differed in posterior areas including the left anterior middle temporal gyrus (MTG). The MTG clusters also extend into superior temporal gyrus (STG), with high frequency having greater activation for both MTG and STG. On the other hand, the low frequency words elicited greater activation for the region that includes left inferior frontal gyrus (IFG) . Generally high activation level for low frequency words was attributed to greater processing resources required for less frequent words. The authors noted that activation found in MTG were coherent with previous findings that MTG is strongly associated with word level processing. However, another fMRI study by Graves et al. (2007) reports a different result. Graves et al. (2007) conducted a picture-naming task and where subjects are presented with a picture and asked to name it. The result indicated higher activation level for low frequency words in the left posterior inferotemporal (IT) cortex, left posterior superior temporal gyrus (pSTG) and the left inferior frontal gyrus (IFG), which was interpreted as being the increased processing cost required in the production of low frequency words. In summary, both studies reported higher activation for low frequency words in IFG, but different results STG or MTG. It is not clear whether the two studies are comparable to each other since they used different types of tasks. In turn, it is not yet completely clear how differing the frequency count affects the cortical processing.

On the other hand, studies using electrophysiological data reveal more on the temporal perspective of word frequency. Laganaro et al. (2009) measured ERP amplitude for words that differ in lexical frequency and cognate status and found that the effect begins to occur around 270∼330ms. Strijkers et al. (2010) found a relatively early effect onset at 170∼200 ms. Both studies used the so-called temporal template proposed by meta-analysis by Indefrey and Levelt (2004) as the basis to map the time windows onto the different functional processes. In turn, Laganaro et al. (2009) posited that relatively later onset falls within the retrieval of phonological form and Strijkers et al. (2010) concluded that the earlier onset of the frequency effect is at the stage of lemma retrieval. Not only are they not consistent to each other, but both conclusions also pose a problem by using a temporal template that assumes strict sequentiality. This assumption potentially overlooks or fails to capture any spatiotemporal dynamics between the processes. Possible interactions between stages makes it almost impossible to estimate which functional process is occuring at particular time point.

Only one prior study looked at both temporal and spatial aspect using intracranial electrophysiology (ICE) (Sahin et al., 2009). The primary focus of the study was to see if there was a sequential order in lexical, phonological, and grammatical processing within Broca's area (IFG region). Sahin et al. (2009) conducted an experiment where subjects either just repeat a written word or insert the word into the blank in a sentence. The blank insertion task had two conditions which required either covertly-inflected (put-put) or overtly-inflected (walk-walked) word forms to make the sentence grammatical. The authors found that lexical information such as word frequency seems to play a role at an early processing stage (0∼200ms) and does not show contrast in following processes like grammatical (200∼320ms) and phonological (320∼450ms) processing in sequence. Not only does this conclusion rest on the somewhat problematic assumption that lexical processing is strictly discrete, it is from the investigation of only one area in the brain, which might have led to missing critical information in other brain regions even though the method is capable of combining the spatiotemporal aspect. The IFG region is known to be responsive to many different functions; however, since there are other parts of brain that are responsive to other processes in production models, it fails to capture the overall time course for each process. In terms of methods, it claimed that the control condition (repetition task) indicated overall weaker phonological or grammatical activation. However, with other tasks that require more explicit involvement of grammatical encoding or phonological encoding, it could have obliterated subtle effects from the word frequency in the plain production condition. If, instead there were only one task (repetition), the effect of word frequency would have become more noticeable if it had turned out to have any effect at all on grammatical or phonological processing.

Summarizing the lines of research, the neuroimaging data has provided empirical evidence on some of the well-attested behavioral effects of the lexical factors in the implementation level. However, since the two sets of information cannot be easily combined, the findings so far have hardly been consistent, making conclusions difficult to draw. It is attributable to the inherent limitations of one neuroimaging technique compared to that of others. To this end, electrocorticogram or electrocorticography (ECoG), introduced in the following section, is expected to clear up some of the inconsistencies and also to offer new evidence on long standing questions concerning the locus of frequency by providing fine-grained images on both spatiotemporal dimensions.

2.3 Electrocorticography and lexical processing

Electrocorticography or ECoG is a technique that bears unique spatial, temporal and functional characteristics. It is an invasive method that requires implantation of subdural electrode on the suface of brain as a part of clinical need in evaluating epilepsy. The proximity of the electrodes to the brain in this invasive method made it possible to study electrophysiological correlates of functional brain activation in more detail than noninvasive recordings (Sinai, 2008). It can provide improved spatial resolution without inteference from the scalp that can play as a low-pass filter. In obtaining the cortical signal, eliminating the filter dramatically increases the specificity of the spatial dimension. Subdural implantation used in ECoG recordings also enables the better detection of event-related activity in a igher gamma (High Gamma: HG) frequency range that extends up to 150 Hz, which has been known to be more accurate and sensitive measure to cortical activation evoked from cognitive tasks like auditory discrimination or word production.

Thus, ECoG has the advantage of incorporating the temporal and spatial resolution at once that no other neuroimaging techniques could have been capable of. fMRI and PET studies can show the participating regions in processing, but does not provide the detailed temporal information, which in turn fails to track the brain activity as it unfolds over time. Electrophysiological methos like EEG or MEG have good temporal resolution but lacks spatial resolution. In short, despite its limitation as to being restricted to a very small number of subjects usually with epileptic brains, ECoG still provides new insights in that it can provide reliable cortical evidence (High Gamma) that occurs within hundreds of milliseconds in functionally specialzed modules in a complex cognitive tasks such as language.

Most of the studies that looked into ECoG recording to study language processing have mainly focused on speech perception so far. Chang et al.(2010) studied the spatio-temporal dynamics underlying simple phoneme detection task where subjects were asked to press button when they heard prespecified sound. It found multiple processes occuring concurrently even with such simple task, implying potentially complex cognitive functions required in excuting the task. Flinker et al. (2011) focused on ECoG recording collected from pSTG region in particular. It found out that auditory functions differ even in the level of sub-centimeter region. One of the main finding was that there were two electrodes adjacent to each other that were still selectively responsive to either phoneme or word. It was indicative of the complex functions within pSTG which have been roughly known as primary auditory cortex as a whole so far.

3 Method

3.1 Subject

One subject (EC2), after a surgical procedure was fitted with a subdural electric grid consisting of 256 electrodes after surgical operation. The grid was placed on the temporal region of the left hemisphere. The subject was male, right-handed and a speaker of standard American English.

3.2 Stimuli

The list used throughout the recording session consisted of 100 English words. The words were relatively evenly distributed when compared against the distribution of words in the dictionary containing 15,289 words (Figure 2). Nevertheless, a majority of the words still belonged to the high familiar category (Nusbaum et al., 1984). Figure 1 shows that on a scale of 1 to 7 where 1 represents an unfamiliar word and 7 a familiar one, almost all the words were a 6 or a 7. The comparison implies that any effect found with the word set can be more safely attributed to the role of frequency, not familiarity.

Figure 1: Word familiarty of the dictionary and the 100 words list

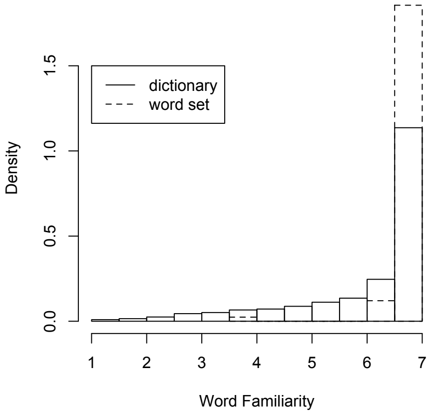
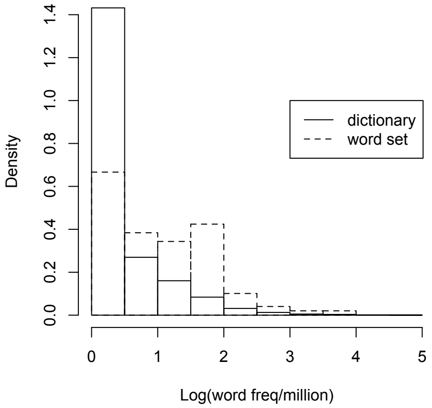


Figure 2: The log frequency value of the dictionary and the 100-word list.



To classify words in terms of their word frequency, the written frequency count from celex was used (Baayen et al., 1995). Written form frequency instead of lemma frequency was selected taking the design of the experiment in which a subject read the written form of the word. [1]One word, ”Thursday,” from the list was missing the frequency value. The remaining words were classified as either High frequency (HF) or Low frequency (LF). If a word’s frequency was greater than the median word’s, it was considered high.

The number of syllables, or syllable length, was also considered in selecting words for analysis. As is shown in Table 1, 80 percent of the high frequency had either one or two syllables. Low frequency words, on the other hand, had a more even distribution in one to three-syllable words. Because the distributional difference in the number of syllables can be a potential confounding factor to the frequency effect, only the subset of one-syllable and two-syllable from high frequency words was selected, which would best match the phonetic properties of the low frequency word counterpart. In other words, out of 25 one-syllable HF word, 12 words were selected and out of 19 two-syllable HF words, 15 words were selected, all of which were matched with one-syllable or two-syllable LF words, respectively.This helped to ensure that any difference in the pattern between two groups was driven by the frequency factor, not by others such as sound complexity.

The phonetic matching process was implemented by calculating phonetic distance between HF and LF words.The scale of phonetic distance is based on a confusability matrix drawn from transcriber disagreements in the Buckeye Corpus (Pitt et al., 2007), which is an indirection of perceptual confusability (Johnson, 2004). Hence, the more confusable the two sounds are, the closer their phonetic distance. As an additional effort to minimize the disparity of the two groups of words in other aspects, only those pairs of words which had a distance of less than 0.62 were selected as words for analysis. Through these processes, out of the high frequency words, total 12 one-syllable words, and 13 two-syllable words were ultimately selected. Table 2 summarizes the mean frequency count and familiarity rating for HF and LF words in one- and two-syllables.

Table 1: The distribution of words by frequency and the number of syllables

Syllable	Low frequency	High frequency
1	12	25
2	15	19
3	15	3
4	6	0
5	2	2
total	50	49

---

[1] Reading a word differs from picture naming in that lexical access occurs via grapheme transcoding at least in intial stage, while in picture naming it starts directly from concept-based lemma acess (Indefrey, 2011). However, the two methods are not considered qualitatively different from each other once it reaches a level of lemma selection as Jacheniak and Levelt (1994) also stated, ”we see no principled reason to expect qualitatively different results with picture stimuli.”

Table 2: Mean frequency and familiarity of the word sets

	one-syllable		two-syllable	
	Mean logFreq [a]	Mean Fam. [b]	Mean logFreq	Mean Fam.
HF	2.12	6.99	1.7	6.93
LF	0.87	6.89	1.08	6.42

[a] CELEX WRITTEN CORPUS

[b] Familiarity rating (1:not familiar, 7:very familiar)

3.3 Procedure

The subject (EC2), a patient in the UCSF Medical Center at Parnassus, sat in a comfortable position and read the list from beginning to end at his own pace. The list had one word per line. The subject repeated the list for four times in one session.[2]

3.4 Data recording

Cortical activity from all 256 channels were recorded at 400 Hz (400 samples/sec) for the entire session.

4 Data analysis

4.1 Data Preprocessing

Following the standard preprocessing procedure used in Chang et al. (2010), all data were first imported into MATLAB for artifact rejection and bad channel filtering. Artifact (discontinuous epochs or abnormal power transient across all frequency bands) was inspected from the raw ECoG signal. Signals from malfunctioning electrodes were also carefully inspected for any abnormal pattern in comparison to the pattern of general activity of other electrodes. After removing the artifacts and abnormal electodes, all channels were referenced to a common averaged reference (CAR, mean of sixteen contiguous channels on the grid of electrodes). notch-filters were then applied to the signals to remove the electrical line noise at 60 Hz, 120 Hz and 180 Hz. The signals were processed through a Hilbert transform to separate the signals at High Gamma (HG) regions (80-150 Hz). Event-related dynamics at HG region was suggested to be more sensitive than signals in lower frequencies in capturing binding roles between cortical regions that are essential in cognitive activity such as speech

---

[2] The subject was recorded for three sessions, making ten iterations of the list in total. However, only the block from the first recording session (four iterations) was used in the study. This was done under the assumption that any potential effect would be greatest for the first session in which the subject had yet to get accustomed to saying the same words repeatedly, which in turn would maximize the difference each group of words invoked.

(Sinai, 2008).

4.2 Signal labelling and Coding

Using the Penn Forced Aligner (Yuan& Liberman, 2008), the audio recording was labelled for the onset and offset of each word. The forced aligner automatically marks the onset and offset of words by matching the acoustic signal to the acoustic models of American English phones. The aligned result was inspected and hand-corrected for more accurate labelling. Each word was then coded as either HF for High frequency or LF as Low frequency, followed by the number indicating syllables in the word. Thus, the words were coded according to one of four categories: HF1, HF2, LF1, and LF2.

4.3 First-pass signal analysis

All event-related cortical signals across all trials and electrode in the High Gamma range (80∼150 Hz) were analyzed during time windows set for one- and two-syllable words respectively (one-syllable word:-600 ms∼570 ms/ two-syllable words: -700 ms ∼720 ms, locked to the onset of word production).

Electrode activiy was normalized to Z-scores. using the mean and stanard deviation of each elecrode during a baseline period at the beginning of the block when the participant was not speaking. The baseline was the longest contiguous silence in which the subjects were presumed to show no event-related activity between 1000 ms∼100000 ms. Normalizing the signal is necessary to make the signals from different electrodes comparable to each other.

4.4 Selection of time window

The length of the time window for the analysis was determined with considering the duration of the pronounced word and the pause that immediately precedes the word. The word production model assumes that much of lexical access occurs even before the subject begins speaking, hence the period of time in which the subject prepared to say the word aloud was selected for analysis. However, the design of the experiment inevitably fails to preclude the possibility that the lexical processing begins from an earlier point (while saying the preceding word aloud and simultaneously looking over the next word). Also, in case that the subject was reading ahead of the list when he was reading the preceding word (the one before the target word), both the signals invoked by the preceding word itself and the one by the target word would occur at the same time. However, it is impossible to tease the two signals apart without knowing the accurate boundary point which signals the precise onset of lexical processing. Given this limitation, the pause between words is regarded as the period where the effect of the other word is minimum and information occuring during lexical processing of the target is the maximum. To this end, the duration of time window was set separately for one- and two-syllable words. The durations of words and pauses for one- and two-syllables and for both high and low frequency words were measured separately so as to capture any possible effect found in behavior pattern and to set the most appropriate time window. The summary of the mean duration for the block is shown in Table 3.

		One-syllable		Two-syllable	
		HF	LF	HF	LF
pause duration (ms)	mean	605	599	713	698
	std	380	221	479	744
word duration (ms)	mean	583	556	693	740
	std	96	85	148	158
Time window for analysis		-600 ms ∼ 600 ms		-700 ms∼ 700 ms	

Table 3: The summary of mean durations and the time window selected

Analysis of variance on the duration as a function of syllable length and word frequency assessed the behavioral pattern. The preliminary result indicated no significant effect of frequency on the duration of the word production ($F = 1.0019$, $p = 0.3187$) and of the pause ($F = 1.8385$, $p = 0.17744$). The number of syllables was a significant factor for the duration of pause ($F = 33.9342$, $p < 0.05$), but not for the word duration ($F = 0.9323$, $p = 0.3360$). There also was an interaction effect between syllable and frequency in the pause duration ($F = 5{,}8306$, $p = 0.02$). Taken together, the results suggest the use of a different size time window for words of different syllable length.

4.5 Selection of electrodes and statistical analysis

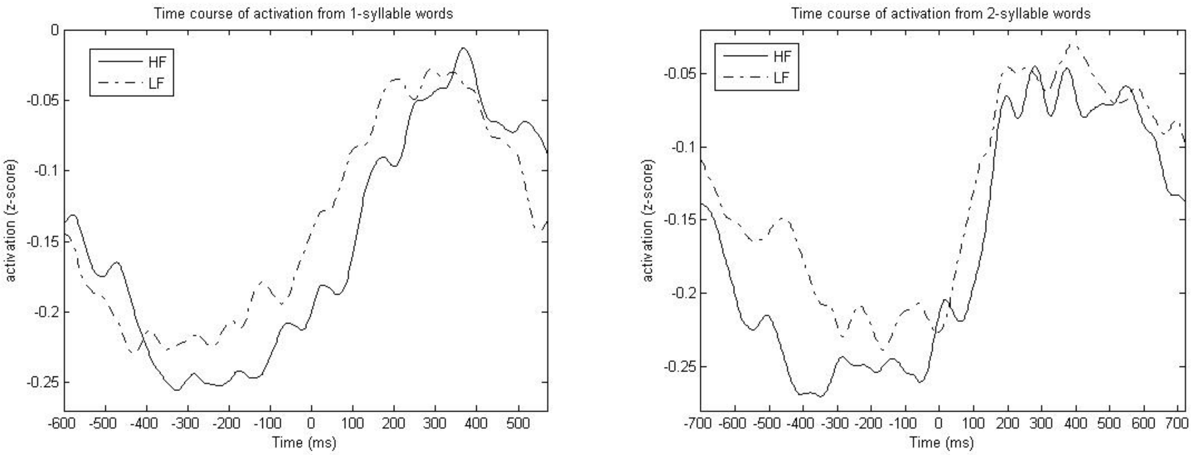
Paired t-tests were performed on all each sub-window that is 100 ms-long intervals (as detailed in the result) to find the location of electrodes that exhibited the greatest difference between HF and LF words and how they gradually changed along the time course. Only the significant differences compared to the baseline mean (paired t-test, $p < 0.01$) were taken into account.

5 Results

5.1 Time course of activation

The ECoG signal was averaged across channeles to give an overall response comparing HF vs LF words. The average resonse is similar to EEG in that localization information is blurred while retaining good temporal resolution. But despite the limitation, it was still useful for observing the overall time course for both one- and two- syllable words, and in turn knowing the general tendency of the activation pattern.

Figure 3 compares the pattern between high and low frequency words in one- and two-syllable words. For one-syllable words, the overall shape of the pattern is very similar for HF and LF words, except that the activation for LF words were almost always higher than that of HF words starting at about -400 ms (-400 ∼300 ms). The activation was higher even after the acoustic onset before it starts to saturate, which suggests that the difference between HF and LF words persisted till the articulation stage. The time point in which the contrast



(a) one-syllable

(b) two-syllable

Figure 3: Time course of activation between high frequency words and low frequency words for one-syllable (left) and 2-syallble words (right).

between HF and LF words reaches its maximum is roughly -150 ms prior to articulation (pre-onset). For two-syllable words, the overall pattern was similar for HF and LF words, but slightly different during pre-onset pause. As with to one-syllable words, LF words showed higher activation than HF words during the pause just prior to production. The contrast is biggest at approximately -450 before the acoustic onset, which is consistent with EEG results that found the effect of frequency about 170 ms before the onset of articulation (Strijkers et al., 2010).

5.2 One-syllable word

In order to observe how the spatial distribution of activated electrodes change as time progresses, the ECoG signals are averaged for sub-time windows at every 100 ms. [3] The pre-onset region (-600 ms∼0 ms for one-syllable words, -700∼0 ms for two-syllable words, 0 ms = acoustic onset) was broken down into sub-windows of 100 ms length. Likewise post-onset or the actual articulation period (0 ms∼ 600 ms for one-syllable words, 0∼700 ms for two-syllable words) was divided to sub-windows of 100 ms length.

The patterns for high and low frequency words were very similar. There were four regions that all showed activation irrespective of frequency and syllable length: superior temporal gyrus (STG), inferior frontal gyrus (IFG), and motor regions (precentral gyrus (PrCG) and postcentral gyrus (PtCG)), and middle tempral gyrus (MTG). Although the function of each region is not yet completely known, a general consensus has been found in previous literature (Indefrey& Levelt ,2004; Indefrey, 2011). Also, broadly considering that selecting lemma is more semantic-oriented and that subsequent processes (mapping the actual sound segments

---

[3] Flinker et al. (2011) found that activation occurs as early as 100 ms after the stimulus onset . Also Indefrey (2011) estimated that each function progresses at around 100 ms interval. This suggests that 100ms is a window that is small enough to capture fine-grained changes, but still not too small in making it easier and clearer to observe the general pattern for spatial distribution of activated electrodes.

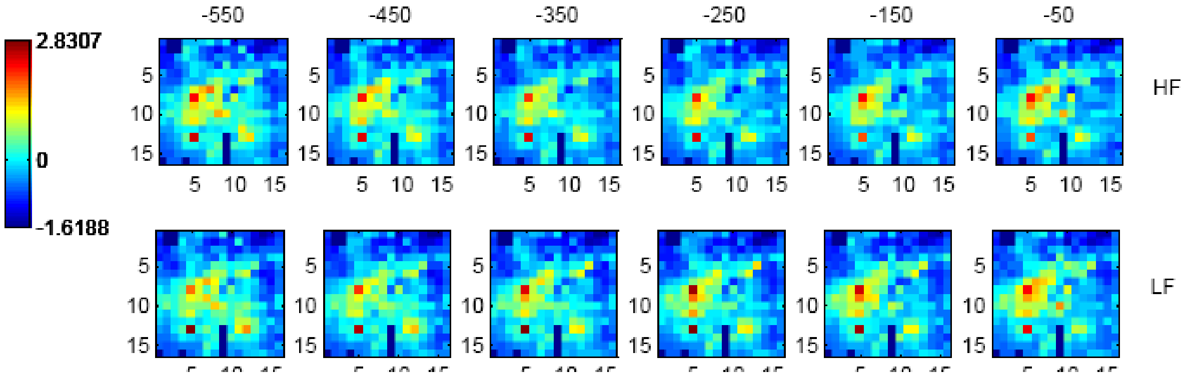
to the lemma) are more phonologically-oriented, the neural signals for the functions are referred as semantic and phonological processing, respectively. The brief summary on the function of each region is presented shortly.

Posterior STG (pSTG) is known to function in the primary auditory perception and respondes selectively to specific speech sounds. In speech production, this region functions in self-monitoring of the outputs of phonological encoding such as syllables. The region is also involved in phonological code (segment) retrieval. Interestingly there is a separate brain region–Inferior Frontal Gyrus (IFG)– that plays a role in syllabification. IFG (also known as Brodmann's 44 or Broca's area) is also found to be a conjunction of complex semantic and phonological processing, under which language comprehension, word analysis, or lexical decision as well as sound-meaning mapping take place. PrCG is a primary premotor cortex where voluntary movement in muscles is planned and excuted in association with pre-motor areas. PtCG is somatosensory motor area. These two regions are related to phonetic encoding and articulation. Lastly, MTG is where most of semantic processing occurs, activated in a process of lemma selection (Indefrey, 2011).

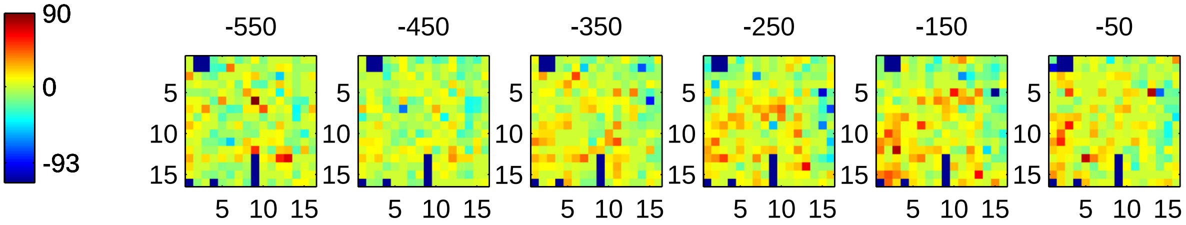
Pre-onset Figure 4a plots the z-scores for high- and low-frequency words. Figure 4a shows that certain regions show stong negatve activation, possibly suggestive of suppression. The electrodes with the positive activation were clustered around the center of the temporal lobe. They were distributed around superior temporal gyrus (STG), and precentral gyrus (PrCG), and postcentural gyrus (PtCG). Part of Middle temporal Gyrus (MTG) adjacent to STG and Inferior Frontal Gyrus (IFG) adjacent to PrCG were also activated but to a lesser degree. Both high and low frequency words showed gradual increase in the activation around the motor region (PrCG, PtCG) as time progresses. The activation is highest in the sixth window, around -50 ms prior to the acoustic onset.

Although the activation of high and low frequency words looks similar, the z-score color map suggests that the activation is higher for LF words within the same eletrode. The difference between conditions (HF vs LF) was assessed for all 256 channels by paired t-tests (Figure 4b). The electrodes that show significant contrasts between HF and LF words become more observable from the third window (-350 ms). The contrast reaches their peak at -150 ms. The region most sensitive to word frequency differences is around IFG, immediately adjacent to the motor regions. There is also a group of electrodes that show a significant difference around -150 ms on the parietal lobe in the vicinity of PtCG. The parietal area is not of central interest to this report and because the region is not where there was noticeable activation in the z-score map aside from few electrodes, they are not further accounted for. Given the grid placement, it might suggest that the orientation of activatinon is to the periphery of central sulcus.

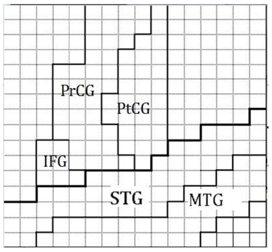
In summary, results show that the most activated and most contrasting regions between HF and LF words are around the anterior STG and IFG, and PrCG. The higher activation level found in LF words in IFG and PrCG indicates that LF words require an extra cognitive load in their phonological processing and motor planning. The relative ease or difficulty in phonological processing could potentially be confounded with other factors like phonotactic probablity. However, since the high and low frequency words were phonetically balanced,



(
t



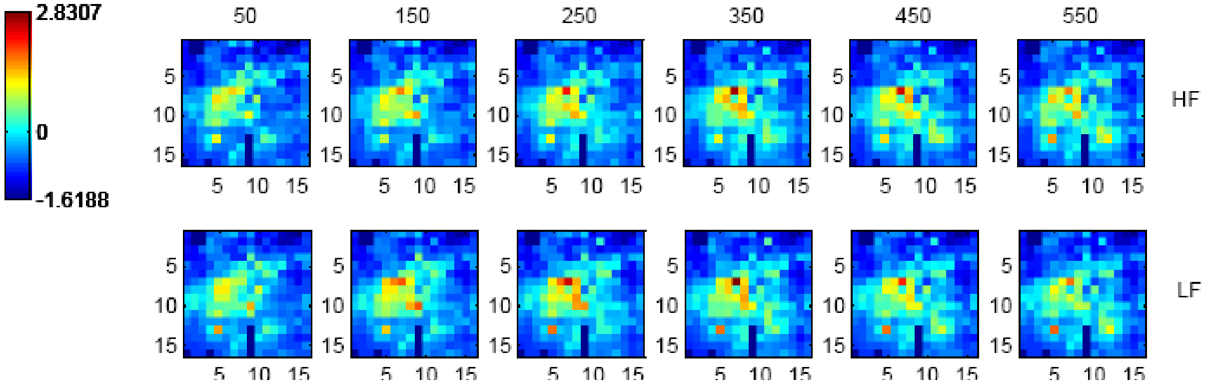
(b) Paired t-test between the HF and LF conditions for each electrode. The number above indicates the center of the 100 ms-window. A positive t value (red color) indicates that low frequency words had higher activation than high frequency words.



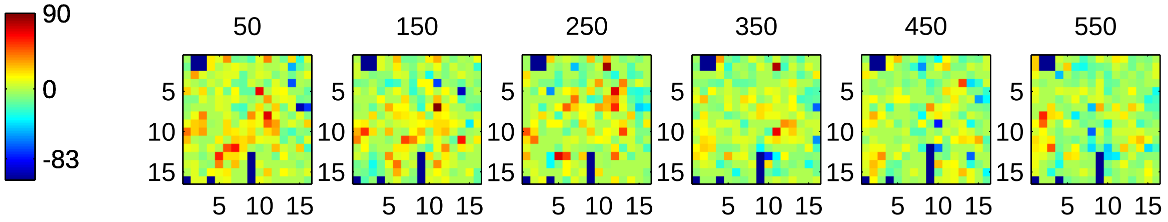
(c) Electrodes placement

Figure 4: Activation of one-syllable words before the acoustic onset of speech (a) Time course of activation during -600 ms∼0 ms. Top: HF words, Bottom: LF words. (b) The result of paired t-test for all electrodes. For each electrode, the mean of LF words was subtracted from the mean of HF words; thus the color scale represents the degree to which LF words differ from HF words (Red: LF >HF; Blue: LF <HF). Only the electrodes that showed statistically significant difference are shown (p <0.01). Not significant electrodes are excluded and shown in light green. (c) The electrode grid is shown in the orientation of the actual grid placement on the cortex. Rough boundaries for their anatomical regions are presented.

the pattern found here might indicate that low frequency words require extra ”effort” in a stage of phonological processing and motor planning (cf. Graves et al. 2007).



(a
fr



(b) Paired t-test between the HF and LF conditions for each electrode. The number above indicates the center of the 100 ms-window. A positive t value (red color) indicates that low frequency words had higher activation than high frequency words.

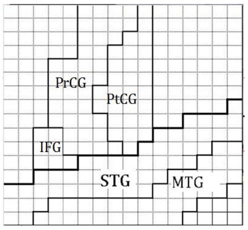
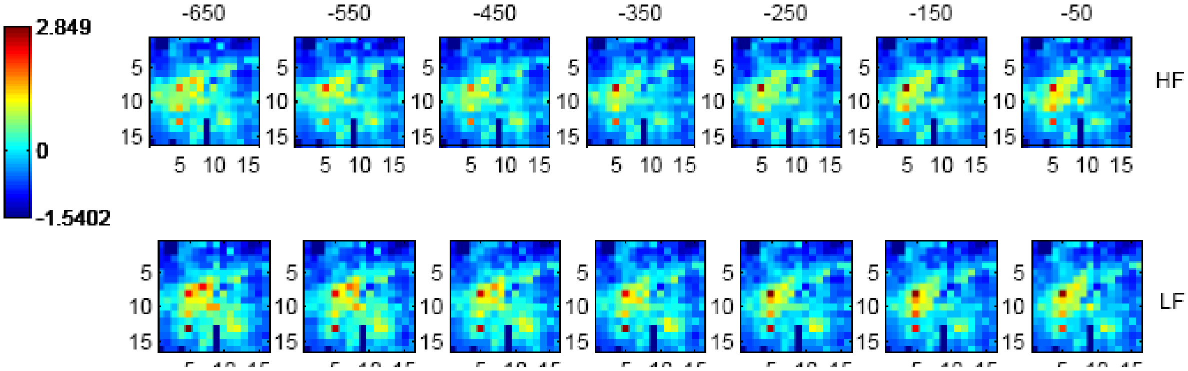
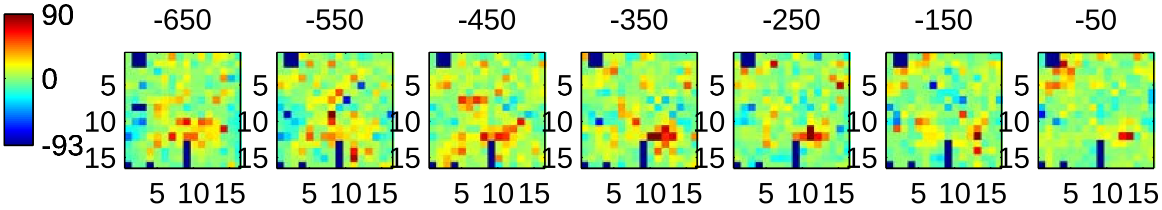


Figure 5: Activation of one-syllable words after the acoustic onset of speech. (a) Time course of activation in 0 ∼ 600 ms. Top: HF words, bottom: LF words. (b) The result of paired t-test for all electrodes. The method and convention are the same as those in Fig. 4

Post-onset The post onset period (0 ms∼600 ms) is shown in Figure 5. The electrodes around PtCG appear to be most active throughout the production for both high and low frequency words. Also, although not of as great amplitude as the signals in the motor region, the electrodes in the STG region began showing a difference between HF and LF around 50 ms after the acoustic onset. Paired t-tests revealed that the difference in the STG region begins to show from around 50 ms throughout 200 ms until it starts to disappear in the fourth window (approximately -450 ms).

The difference in the motor regions peaks in the third window (approximately 200 ms∼300 ms after the onset) before beginning to decrease. The differences in signal are in line with





(b) Paired t-test between the HF and LF conditions for each electrode. The number above indicates the center of the 100 ms-window. A positive t value (red color) indicates that low frequency words had higher activation than high frequency words.

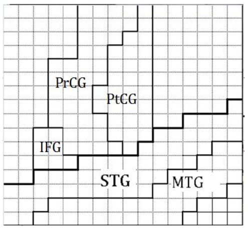


Figure 6: Activation of two-syllable words before the acoustic onset of speech. (a) Time course of activation from -700 ms to 0 ms. Top: HF words, bottom: LF words. (b) The result of a paired t-test from all electrodes. The method and convention are the same as those used in Fig. 4.

what is shown in the time course graph in Figure 3. These results show that activations in brain region peak around midway through word production and begin to saturate after the peak.

The activation in the STG region after the acoustic onset might reflect self-listening. It is noteworthy to see that LF words invoked a higher activation level even in phonological processing in perception. Such a pattern is consistent with the signal pattern found in the IFG region at pre-onset, suggesting that it requires a greater load to assemble and process the LF words phonologically than it does HF words.

5.3 Two-syllable word

Pre-onset As two-syllable words were longer in terms of their duration (-700∼700ms, aligned at the acoustic onset at 0 ms), they were divided to seven sub-windows for both before and after the acoustic onset.

The overall pattern of two-syllable words during pre-onset is shown in Figure 6. The pattern is simiar to one-syllable words in that both STG and PrCG regions are highly activated for either one or two syllable words with generally higher activation for LF words. In terms of the specific temporal progression over time shown in Figure 6a, the activation was highest in the second and third window, around 500 ms prior to the onset. The activation gradually decreased until the acoustic onset. This pattern is more noticeable with LF words than with HF words, suggesting that LF words were indeed more activated than HF words. The electordes around MTG, posterior STG (pSTG) showed some activation during the pre-onset period, although they are relatively weak compared to the signal from the motor region.

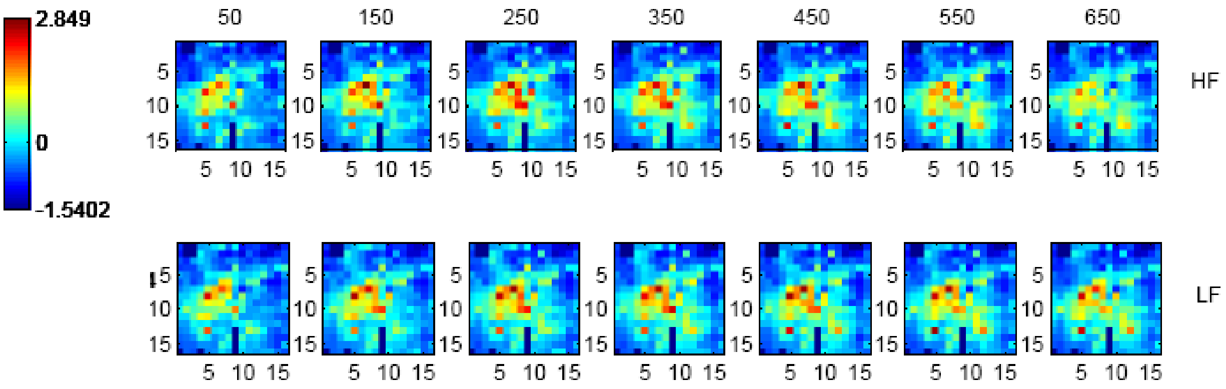
The result of the paired t-test comparing the processing of LF and HF words is shown in Figure 6b. It shows that the difference in the magnitude of activation in PrCG and STG seems to occur as early as 500 ms prior to onset. The contrast in the activation level for both HF and LF condition is more noticeable in MTG and pSTG than those in the motor regions around 400 ms to 200 ms prior to the onset (the 4th and the 5th time-window). It seems that the difference is around the central region near the sylvian fissure and shifts to the more peripheral regions as time progresses (rougly from STG to MTG).

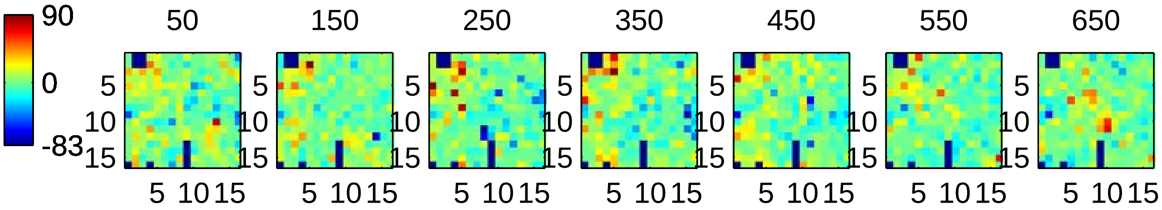
Post-onset The post-onset region was also sub-divided into 7 sub-windows. The activation pattern shown in Figure 7was generally similar to the post-onset period of one-syllable words. The electrodes around the PrCG/PtCG remain highly activated throughout the time course for both LF and HF words, with LF words occasionally showing slightly higher activation. The activation peaks around 350 ms post onset, which is almost half the total duration of the production. This pattern is consistent with the result of the post-onset activation from one-syllable words.

The result of paired t-tests shows that little difference is actually caused by the frequency condition. It suggests that both HF and LF words caused very strong activation in the production period, possibly due to the voluntary movement of articulators while reading. After the activation reaches its peak and begins to dicrease midway through the entire duration, the activation for LF words seems to saturate at a slightly faster rate than that for HF words, as is supported with more negative values from the 4th window (480 ms∼).

6 Summary and Discussion

The aim of this study was to observe the spatiotemporal distribution of cortical activity as a function of word frequency so as to shed light on how neural processing of high and low frequency words may differ. One-syllable words and two-syllable words were observed separately for possible difference in dynamics of each sub-process in speech production and





(b) Paired t-test between the HF and LF conditions for each electrode. The number above indicates the center of the 100 ms-window. A positive t value (red color) indicates that low frequency words had higher activation than high frequency words.

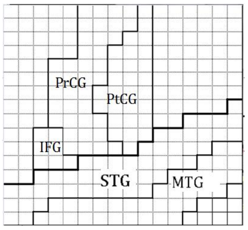


Figure 7: Activation of two-syllable words after the acoustic onset of speech. (a) Time course of activation in 0 ∼700 msec. Top: HF words, bottom: LF words. (b) The result of a paired t-test from all electrodes. The method and convention are the same as those used in Fig 4.

how they are modulated by word frequency. The subject saw a list of 100 English words of different frequency counts and said each word one at a time. The ECoG recording was time-aligned to the acoustic onset of word production, where the subject began speaking the word. The period prior to speech onset as well as the period of actual production was analyzed.

High and low frequency words showed a very similar pattern overall. Several brain regions showed increased activation during the spech task encompassing STG, IFG PrCG/PtCG, and MTG regions, while other regions showed suppression of strong acitivation. ()During the pre-onset period, a certain level of high gamma activation in IFG, PrCG, and MTG is observed. PrCG/PtCG region and pSTG are more activated during post-onset period. This pattern of neural firing suggests that the reading-aloud task in this study required combinations of such complex processing as visual perception, speech production (which also included speech perception to some extent), thus essentially lexical access. Among these regions, the area most highly activated throughout the time course was the motor regions (PrCG/PtCG), which reflects the fact that the subject was actively involved in the articulation. Although the experimental setting did not explicitly include a perception task, pSTG shows slight activation 300 msec prior to articulation for two-syllable words and during articulation for both one and two syllable conditions. The activation in pSTG supports the idea that this regions is activated both for sound perception and production (Peramunage et al., 2011).

Processing of high and low frequency words was found to be different at the neural level, which can shed light on the locus of word frequency and speech production models. Firstly, the contrasts found in PrCG and PtCG regions support the idea that word frequency plays a role in the phonetic encoding process, contra Jescheniak and Levelt (1994). The contrast appears to occur primarily before the subject starts to speak, but less so during actual articulation. Such a pattern is in line with the finding that the activation occurs at a stage as early as 200 ms after the stimulus presentation (Sahin et al., 2009). Sahin et al. (2009) claimed that it was at this early stage of processing where lexical information like word frequency is incorporated. Motor regions might be expected to differ for HF and LF words during the production stage also, as it is possible that the durational difference found for HF and LF words has some correlate at the neural level as well. However, the motor regions did not show much difference during the actual articulation. The contradiction might be attributable to the behavioral result found where the durational difference was not significant. That is, unlike the results from spontaneous speech (cf. Gahl, 2008), the procedure of the experiment was not sufficient in eliciting the natural reduction process found in behavioral experiments, such that little contrast was yielded for articulation.

Secondly, the neural activation was also different in regions that function in semantic and phonological processing. Low frequency words exhibited overall greater High Gamma activation in the regions important for speech production models, IFG, MTG, and pSTG. Of those, MTG and pSTG have been identified to represent the processing of meaning (MTG) and sound (pSTG), respectively (cf. Indefrey & Levelt, 2011). The result found in the current study indicates that MTG shows a significantly high activation for LF words during pre-onset and early post-onset period (0 ms$\sim$200ms in two-syllable words). Also, the electrodes in pSTG indicated greater level of activation for LF words when the subject was presumably perceiving the sound. Unlike MTG and pSTG, the role of IFG is relatively unclear. Poldrack et al., (1999) earlier found that IFG is activated for both phonological and semantic processing during reading. More recent evidences confirmed that IFG is involved with the activity related with both semantic and phonological processing, although it is primarily associated with syllabification (cf. Indefrey & Levelt, 2004; Indefrey, 2011). Hence, the difference in the activation level found in those three areas (pSTG, MTG, and IFG) is likely to represent a greater cognitive load for low frequency words in accessing both meaning and the phonological form. In sum, low frequency words appear to require greater endeavor in order to map the correct semantic referent and to the phonological form, and to the actual production.

A final point to note is that the word frequency effect differs for one- and two-syllable words. Figure 3 suggests that, although low frequency words evoke a higher amplitude high gamma signal than high frequency words do, the activation of HF and LF words for two syllable words begins to diverge 200 to 300 msec earlier than it does for the one syllable words. The difference in signal amplitude between HF and LF words is greatest between -200 ms and -100 ms prior to the acoustic onset for one-syllable words and between -500 ms and -400 ms for two-syllable words. The spatial configuration of the 256 electrodes at each 100ms-long-time window points to the different brain regions where the contrast is most noticeable. A closer look from the spatial dimension reveals that different regions of the brain are highlighted for the contrast between HF and LF words for one- and two-syllable words. Specifically, the graph of the spatial distribution shows that the difference is at IFG and motor regions (PrCG/PtCG) for one-syllable words, whereas it is at MTG, premotor region (PrCG), and pSTG for two syllable words during 500 to 400 msec pre-onset. Not only the locations, but the magnitude of contrast also differed, as it was greater for two-syllable words.

It is surprising to see that some regions are selectively involved with frequency depending on the syllable length. Firstly, MTG, the region that primarily functions as semantic processing, shows contrast for frequency for two-syllable words only. In addition, the effect of frequency was shown in the IFG region for one-syllable word only but not in two-syllable words. Secondly, as many researchers have found, IFG is primarily engaged in syllabification along with other functions (Sahin et al., 2009; Indefrey and Levelt, 2004; Indefrey, 2011). It is interesting to see that there is no difference due to frequency in IFG for two-syllable words, in which more segments need to be syllabified than in one-syllable words. Instead, a greater degree of contrast is observed in pSTG (the region related to sound processing). Previous studies on word production have suggested that pSTG is associated with the load in self-monitoring of what is to be produced (outputs from phonological encoding process) (cf. Indefrey and Levelt, 2004). Thus, it is implied that the cognitive load required during self-monitoring is the effort to assemble the syllables in sequence.

To this end, the result sheds light on the two theoretical concepts from the psycholinguistic literature. Firstly, no processing difference in IFG for two-syllable words and a greater difference found in pSTG for one-syllable words raise the possibility of mental syllabary, a concept that assumes speakers to retrieve pre-stored commands for syllable articulation after phonological encoding is done (Levelt et al., 1999). If the mental syllabary exists, little difference is expected in syllabification process between one- and two-syllable words, since the subject could have just retrieved the frequent syllables without drastic increase from online computational load in syllabifying the string of segments.

The other concept concerns the relations between processes in the production model: cascading and interactions between each functional stage (cf. Griffin & Ferreira ,2006). Assuming that pSTG is associated with self-monitoring and syllable assembly, less labor would be required in the syllable assembly for one-syllable words, which would lead to the less load for self-monitoring. On the other hand, two syllable words would require greater load in self-monitoring, which in turn would have possibly bottlenecked the flow of cortical downstream activities throughout the production processes. As a result, other activities of relatively less processing effort would have been obliterated by the bottleneck. This postulation is only sensible when continuous flow of activation and interaction between stages are assumed. Cascading of activation is further evidenced by contrasts in the PrCG regions, since the processing difference found in this region indicates that lexical property can also affect articulatory processes. The finding that motor regions are also modulated by the lexical factors (the evidence of cascading activation) is consistent with the similar study on the neural substrates of neighborhood density (Peramunage et al., 2011).

The present result suggests that semantic and monitoring process of longer words is different from shorter words, contra the finding by Sahin et al. (2009) who reported no sensitivity to word length. The result also implies that words of different syllable length do not always have identical processes during production, which might explain the inconsistent results in Balota and Chumbley (1985) and Jescheniak and Levelt (1994).

7 Potential Problems and Conclusion

Several caveats should be mentioned. The most noticeable problem is the lack of clarity regarding the onset of lexical activation. The problem is strongly associated with the inherent characteristics of the task where the only consistently observable point in each trial was the onset of articulation. However, it is likely that most of the lexical processing had already been completed by that moment. The problem is complicated by the method of stimulus presentation during the recording, where the subject read the list at his own pace and he could look ahead on the list. To alleviate the problem, time windows of different length (600 ms for one-syllable words; 700 ms for two-syllable words) were used for analysis under the assumption that the more robust lexical activation would begin only after the subject paid active attention to the target word. To confirm the assumption, a different time alignment is necessary, possibly to the offset of the preceding word (i.e., the onset of the pause).

Related to the unclear onset of lexical access, an additional treatment of the data is also necessary to eliminate the possible effects from the previous or following words. One such treatment would be to take into account the frequency of adjacent words. It might be the case that a HF word followed by another HF word differs from one followed by a LF word, or vice versa. Hence, looking at the frequency of the word immediately preceding or following a certain word might help delineate the confounding effects.

The effect of repetition is also potentially problematic. The study used only the first four iterations where the processing difference caused by frequency was most observable, taking the other two later blocks out in the analysis. Still, having the subject repeat the same list for multiple times could have saturated any existing effect during the recording. In order to tell if there is any missing information blurred by the repetiton, a more fine-grained temporal account within the four iterations would be necessary.

Another thing to note is the difficulty in generalizing the result. For instance, only one- and two- syllable words were used in the study, which made it hard to generalize about what the cognitive load associated with assembling syllables is like. If more multisyllabic words are included, the effect of syllable length can be clarified more accurately. Also, considering only word frequency is another potential factor that makes the generalization hard, especially when word frequency can interact with other lexical properties (Weekes, 1997; Yap & Balota, 2008; Hauk, 2008). Adding more factors into the analysis will contribute to removing any confounding effects from other sources and in turn will assume that the result is solely attributable to frequency. In particular phonotactic probablity and syllable frequency is relavent for the result in this study in that the less phonotactically probable words should require greater effort in phonologocal processing and considering the idea of syllabary, the more frequent syllables should require less effort. Coding the data taking such factors into account would potentially capture what was not revealed with the frequency effect alone. One fundamental solution to insufficient generalizability is to increase the data size. Increasing the number of words in the analysis will enable in clarifying some of the inconclusive findings.

Despite the methodological concerns, this study opens a new empirical ground on how word frequency affects processing at the neural level. The result showed that multiple aspects of neural processing are affected by word frequency. The comparison between one-and two-syllable words sheds light on the underlying reasons for increased processing cost for longer words. It was found that certain brain regions are sensitive to syllable length and that the MTG region (semantic processing) participates more for processing longer words, which increased the processing load overall.

References

Baayen, R. H., Piepenbrock, R., and Gulikers, L. (1995). The CELEX lexical database [CD-ROM].

Balota, D. A. and Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language*, 24(1):89–106.

Canolty, R. T., Soltani, M., Dalal, S. S., Edwards, E., Dronkers, N. F., Nagarajan, S. S., Kirsch, H. E., Barbaro, N. M., and Knight, R. T. (2007). Spatiotemporal dynamics of word processing in the human brain. *Frontiers in neuroscience*, 1(1):185–196.

Carreiras, M., Mechelli, A., and Price, C. J. (2006). Effect of word and syllable frequency on activation during lexical decision and reading aloud. *Hum. Brain Mapp.*, 27(12):963–972.

Chang, E. F., Edwards, E., Nagarajan, S. S., Fogelson, N., Dalal, S. S., Canolty, R. T., Kirsch, H. E., Barbaro, N. M., and Knight, R. T. (2010). Cortical spatio-temporal dynamics underlying phonological target detection in humans. *Journal of Cognitive Neuroscience*, 23(6):1437–1446.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological review*, 93(3):283–321.

Flinker, A., Chang, E. F., Barbaro, N. M., Berger, M. S., and Knight, R. T. (2011). Sub-centimeter language organization in the human temporal lobe. *Brain and Language*, 117(3):103–109.

Forster, K. and Chambers, S. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, 12(6):627–635.

Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, 84(3):474–496.

Graves, W. W., Grabowski, T. J., Mehta, S., and Gordon, J. K. (2007). A neural signature of phonological access: distinguishing the effects of word frequency from familiarity and length in overt picture naming. *Journal of cognitive neuroscience*, 19(4):617–631.

Griffin, Z. and Ferreira, V. S. (2006). *Properties of spoken language production*, pages 21–59. Elsevier.

Hauk, O. and Pulvermüller, F. (2004). Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology*, 115(5):1090–1103.

Indefrey, P. (2011). The spatial and temporal signatures of word production components: a critical update. *Frontiers in psychology*, 2.

Indefrey, P. and Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, 92(1-2):101–144.

Jescheniak, J. D. and Levelt, W. J. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4):824–843.

Johnson, K. (2004). Aligning phonetic transcriptions with their citation forms. In *Acoustics Research Letters Online*, number 5, pages 19–24.

Kittredge, A. K., Dell, G. S., Verkuilen, J., and Schwartz, M. F. (2008). Where is the effect of frequency in word production? insights from aphasic picture-naming errors. *Cognitive neuropsychology*, 25(4):463–492.

Laganaro, M., Morand, S., Schwitter, V., Zimmermann, C., Camen, C., and Schnider, A. (2009). Electrophysiological correlates of different anomic patterns in comparison with normal word production. *Cortex*, 45(6):697–707.

Levelt, W. J., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *The Behavioral and brain sciences*, 22(1).

Monsell, S. (1991). *The nature and locus of word frequency effects in reading*, chapter 6, pages 148–197. Routledge.

Nusbaum, H. C., Pisoni, D. B., and Davis, C. K. (1984). Sizing up the hoosier mental lexicon: Measuring the familiarity of 20,000 words. Technical report.

Oldfield, R. C. and Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, 17(4):273–281.

Peramunage, D., Blumstein, S. E., Myers, E. B., Goldrick, M., and Baese-Berk, M. (2011). Phonological neighborhood effects in spoken word production: an fMRI study. *Journal of cognitive neuroscience*, 23(3):593–603.

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., and Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *NeuroImage*, 10(1):15–35.

Prabhakaran, R., Blumstein, S. E., Myers, E. B., Hutchison, E., and Britton, B. (2006). An event-related fMRI investigation of phonological–lexical competition. *Neuropsychologia*, 44(12):2209–2221.

Sahin, N. T., Pinker, S., Cash, S. S., Schomer, D., and Halgren, E. (2009). Sequential processing of lexical, grammatical, and phonological information within broca's area. *Science*, 326(5951):445–449.

Savin, H. B. (1963). WordFrequency effect and errors in the perception of speech. 35(2):200–206.

Sinai, A. and Sinai, A. (2008). Brain mapping of language using high frequency gamma oscillations in the electrocorticography brain research in language. In Breznitz, Z., Joshi, R. M., Joshi, R. M., and Breznitz, Z., editors, *Brain Research in Language*, volume 1 of *Literacy Studies*, chapter 10, pages 247–273. Springer US, Boston, MA.

Strijkers, K., Costa, A., and Thierry, G. (2010). Tracking lexical access in speech production: Electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex*, 20(4):912–928.

Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *The Quarterly Journal of Experimental Psychology Section A*, 50(2):439–456.

Whalen, D. H. (1991). Infrequent words are longer in duration than frequent words. 90(4):2311.

Yap, M. J. and Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language*, 60(4):502–529.

Yuan, J. and Liberman, M. (2008). Speaker identification on the SCOTUS corpus.