Available online at www.sciencedirect.com
SCIENCE DIRECT®
Brain and Language 90 (2004) 88–94
Brain and Language
www.elsevier.com/locate/b&l

The precise time course of lexical activation: MEG measurements of the effects of frequency, probability, and density in lexical decision

Linnaea Stockall,* Andrew Stringfellow, and Alec Marantz

Department of Linguistics and Philosophy, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Accepted 2 December 2003
Available online 13 February 2004

Abstract

Visually presented letter strings consistently yield three MEG response components: the M170, associated with letter-string processing (Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin, 1999); the M250, affected by phonotactic probability, (Pylkkänen, Stringfellow, & Marantz, 2002); and the M350, responsive to lexical frequency (Embick, Hackl, Schaeffer, Kelepir, & Marantz, 2001). Pylkkänen et al. found evidence that the M350 reflects lexical activation prior to competition among phonologically similar words. We investigate the effects of lexical and sublexical frequency and neighborhood density on the M250 and M350 through orthogonal manipulation of phonotactic probability, density, and frequency. The results confirm that probability but not density affects the latency of the M250 and M350; however, an interaction between probability and density on M350 latencies suggests an earlier influence of neighborhoods than previously reported.

Keywords: MEG; Lexical decision; Lexical access; Lexical frequency; Phonotactic probability; Neighborhood density effects; M350

1. Introduction

Previous MEG studies of visual word recognition have identified three response components that appear consistently across subjects and across conditions (Embick, Hackl, Schaeffer, Kelepir, & Marantz, (2001); Helenius, Salmelin, Service, & Connolly, 1998, 1999; Koyama, Kakigi, Hoshiyama, & Kitamura, 1998; Kuriki et al., 1998; Kuriki, Takauchi, Fujimaki, & Kobayashi, 1996; Pylkkänen, Stringfellow, Flagg, & Marantz, 2000; Pylkkänen et al., 2002; Sekiguchi, Koyama, & Kakigi, 2000). As shown in Fig. 1, the M170 is associated with a bilateral field distribution over the occipitotemporal sensors;[1] the M250 with a left-lateralized dipolar pattern oriented along the lateral axis, displaying a posterior positive field and an anterior negative field, and the M350 with a left-lateralized dipolar distribution oriented along the anterior-posterior axis, showing the positive field on the right and the negative field on the left.

Tarkiainen and (Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin (1999) found evidence that in the left hemisphere, the early occipito-temporal activity (M170) reflects pre-lexical visual processing that is specific to letter-strings while the right hemisphere activity is associated with aspects of visual object processing common to both letters and symbols.

M350 latencies and/or amplitudes have been shown to parallel RTs in being sensitive to repetition (Pylkkänen, Stingfellow, & Marantz, 2002 Sekiguchi et al., 2000), frequency (Embick et al., 2001) and cloze probability (Helenius et al., 1998, 1999). See Pylkkänen & Marantz (2003) for a discussion about the connection between M350 and ERP N400.

Pylkkänen et al. (2002) investigated the factors affecting the timing of the M350 by manipulating the relative phonotactic probability and phonological neighborhood density of visually presented letter stings in a lexical decision experiment. Phonotactic probability is the probability of occurrences of individual phonemes and the probability of pairs or sequences of phonemes.

* Corresponding author. Fax: 1-617-253-5017.
E-mail address: l-stock@mit.edu (L. Stockall).

[1] The M170 most likely originates from two separate sources, one in each hemisphere (see Tarkiainen et al., 1999) for discussion of the functional specializations of the two sources.

0093-934X/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S0093-934X(03)00422-X

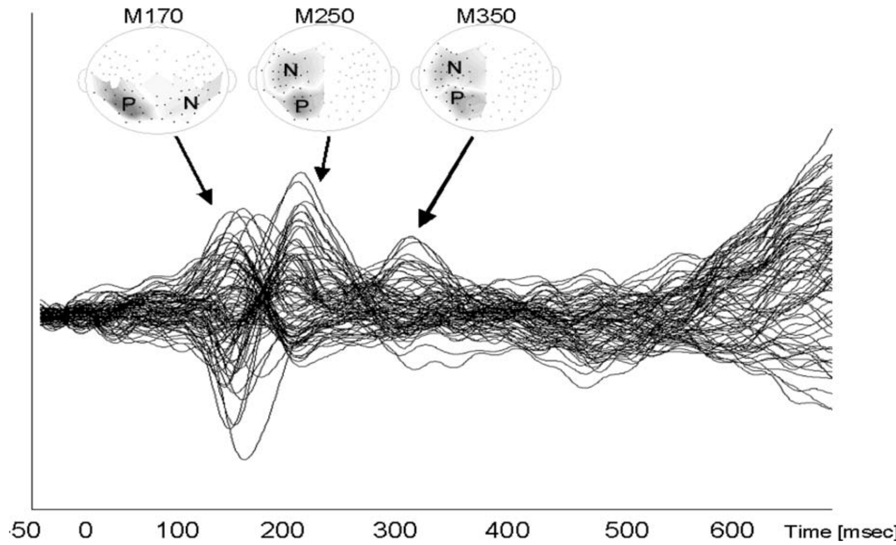

Fig. 1. The magnetic field distributions of the M170, M250, and M350 response components at the time of component peak in one representative participant. The letter P indicates the positive field (i.e., the magnetic field emerging from the brain) and the letter N the negative field (i.e., the magnetic field entering the brain) of the magnetic field around the current source.

Phoneme frequency depends on the number of words in a language that contain the phoneme and the frequency of occurrence of those words. The probability of phoneme pairs reflects the sequential probabilities of phonemes of words in a language. The neighborhood density of a word is a measure of the number and frequency of similar words in a language. In the current experiment, two words are considered to be in the same neighborhood if they can be related by the substitution of one phoneme or letter[2] (depending on whether phonological or orthographic density is at issue). For example, *cat*, *jar*, and *cur* are all neighbors of *car*. Neighborhood density and phonotactic probability are highly correlated in natural language—words whose constituent bi-phones are very common in the language tend to be similar to a large number of other words in the language. Pylkkänen et al. (2002) found that stimuli that are high in both phonotactic probability and neighborhood density elicited longer response times in lexical decision relative to stimuli with lower probabilities and neighborhood densities. This result is consistent with the findings of Vitevitch & Luce (1999). The inhibitory effect of neighborhood density on lexical decision times can be understood as being due to the fact that high density items necessarily resemble, and hence activate, more actual lexical entries than low density items, and the more competing lexical entries an item activates, the longer the time needed for determining whether one of them can be selected as the "winner."

The key finding of Pylkkänen et al. (2002) was that the same high probability, high density items which elicited delayed RTs were associated with earlier M350 latencies (facilitation) than the low probability/density stimuli. The M350 component is, then, not only the first component sensitive to factors affecting lexical activation (repetition, frequency, and cloze probability), but seems to index a stage of activation prior to processes of competition and selection. If the M350 were sensitive to competition among activated lexical items, it should show inhibition for the high density items rather than facilitation.

Pylkkänen et al. also found that high probability/density stimuli were correlated with decreased M250 amplitudes, suggesting that sub-lexical frequency information affects earlier stages of processing than lexical frequency. Decreased amplitude of an evoked brain response constitutes a facilitatory effect. The later facilitation indexed by the latency of the M350 response is then plausibly a consequence of the earlier M250 facilitation.

Luce & Large (2001) manipulated phonotactic probability and neighborhood density orthogonally to explicitly confirm the conclusions of Vitevitch & Luce (1999) that density inhibits and probability facilitates lexical decision. The current experiment expands on the design of Luce and Large, with the addition of lexical frequency as a stimulus variable, to answer two questions. The first question was whether phonotactic probability alone (unconfounded from density) would facilitate the processing indexed by M250 and M350 components, while density alone would have no effect on either of these stages. This result would confirm that both M250 and M350 index precompetition stages of lexical processing. The second question concerned possible interactions between sub-lexical and lexical level frequency in cases where those variables do not correlate.

[2] Addition and subtraction of a segment are also often allowed in defining neighborhoods (see for example Vitevitch & Luce, 1999).

2. Materials and methods

2.1. Participants

Twenty-five right-handed, English-speaking adults with normal or corrected-to-normal vision gave their informed consent to participate in the experiment (thirteen females and twelve males ranging in age from 18 to 32, mean age 23.5). Participants were all students or employees at the Massachusetts Institute of Technology and were paid $20 for their participation.

2.2. Stimuli

A total of 320 word stimuli and 160 nonword stimuli were prepared. All words were uninflected and were not homophonous with other words. For the words, three factors were explicitly manipulated: orthographic frequency, biphone probability, and phonological neighborhood density. For the nonwords, only biphone probability and phonological density were manipulated. Stimuli with different values for a factor comprised non-overlapping distributions. For example, all the low density items had lower neighborhood density counts than any of the high density items.

Orthographic frequency was obtained from the MRC Psycholinguistic Database and is based on Kucera & Francis (1967). Low frequency words had fewer than 10 occurrences per million (average of 3.6, range from 0[3] to 10), and high frequency words had a frequency of occurrence greater than 10 per million (average of 46.5, ranges from 11 to 228).

Biphone probability was defined as the weighted frequency (per Kucera & Francis, 1967) of the biphones (CV and VC) that comprised the words.[4] Low biphone probability items were defined as having an average biphone probability of less than 300 (Words: average of 179; range from 28 to 298; Nonwords: average of 179; range from 2 to 297). High biphone probability items had an average biphone probability of greater than 300 (Words: average of 719; range from 316 to 2837; Nonwords: average of 781; range from 302 to 4788).

Phonological density was defined as the phonological neighborhood of an item computed as the number of words formed by appropriate substitution of a phoneme at any position. An inventory of 15 vowels and 24 consonants was used.[5] High density items had a minimum of 17 neighbors (Words: average of 22.7, range of 17–35; Nonword: average of 22.5, range of 17–35). Low density items had a maximum of 16 neighbors (Words: average of 12.4, range of 4–16; Nonwords: average of 12.3, range of 2–16).

All test items had a CVC syllable structure, which can be realized in English by as few as 3 letters (*dog*) or as many as 7 (*thought*). The words used in this experiment were all mono-morphemic and the nonwords used as test items were limited to those with regular orthography to phonology mappings (this was determined by having three native English speakers read possible nonword stimuli aloud—items on which there was less than perfect agreement among all native speakers were excluded). Obeying these restrictions lead to a maximum length of 5 letters for the test items. Two ANOVAs comparing item length were performed, one each for words and nonwords. There were no significant differences in item length across stimulus conditions. An additional 160 filler nonwords (matched for length with the test nonwords) were created. These items were not included in any analyses.

2.3. Procedure

Stimuli were presented using PsyScope 1.2.5 (Cohen, MacWhinney, Flatt, & Provost, 1993) in a randomized order. Each trial consisted of a fixation point (+) that lasted for 1000 ms followed by the presentation of the stimulus, which disappeared at the button press response. The task was continuous lexical decision. Participants used their left index and middle fingers to press the response buttons (the left hand was used in order to minimize the amount of left hemisphere activity associated with motor control). The intertrial interval randomly varied between 100 and 900 ms.

Neuromagnetic fields were recorded using an axial gradiometer whole-head 93 channel system (Kanazawa Institute of Technology, Japan). Data were sampled at 1000 Hz, with acquisition between DC and 200 Hz. The recording for each participant lasted approximately 25 min. External noise sources were removed from the data using the continuously adjusted least-squares method (CALM, Adachi, Shimogawara, Higuchi, Haruta, & Ochiai, 2001). Responses to stimuli were averaged by stimulus condition. In the averaging, artifact rejection was performed by excluding all responses to stimuli that contained signals exceeding ± 2.5 pT in

[3] Several words with "0" frequency were included; these were not unusual words, e.g., "chime" and "bike."

[4] This measure was obtained for all CVC items in English. To compute this value, all CVC entries in the Carnegie Mellon Pronouncing Dictionary (v. 0.6, 1998) were found. For items with multiple pronunciations, only the first pronunciation entry per the Merriam Webster Online Dictionary (2002) was considered. A total of 1502 CVC words, with 1320 unique pronunciations, were considered. Items which comprised well-formed English words but which had 0 occurrences per million were tabulated as having a frequency of 0.1. Based on this list, the frequency of each biphone was determined.

[5] In English there are 22 possible word-initial consonants /ŋ/ and /ʒ/ are not possible word-initial phonemes, 15 possible vowels (/ɚ/ was considered a vowel), and 21 possible word-final consonants (/j/, /w/, and /h/ are not possible word-final phonemes—glides that formed part of dipthongs were not counted as consonants).

amplitude. Epochs were also excluded from further analysis based on reaction time criteria. Following averaging, data were baseline adjusted using a 100 ms prestimulus interval and low pass filtered under 30 Hz.

2.4. Data analysis

Reaction times were calculated from the onset of the visual stimulus. Incorrect trials and RTs deviating over 3 SD from the mean for the particular participant were excluded from the analysis. This resulted in the exclusion of 7.7% of the data. These trials were also rejected from the MEG averages. Subjects with an overall error rate of higher than 10% were rejected from further analysis. The data from five participants did not survive this criterion, leaving 20 subjects whose brain data was analyzed. Only MEG averages consisting of more than 30 trials after artifact and error rejection were accepted for further analysis.

In the analysis of the MEG data, averaged signals were first visually inspected to identify dipolar field distributions that showed consistency across experimental conditions and across participants. Since the aim of the present study was to investigate the effects of the stimulus variables on the timing of the M250 and M350, subjects for whom these response components were not identifiable in all conditions were not considered in the analysis (four subjects were rejected for this reason).

The amplitudes and latencies of the three evoked response components (the M170, M250, and M350) were recorded by first picking the best two sensors for each component (the left hemisphere sensors that had the greatest ingoing and outgoing amplitude). The root mean square (RMS) field strength from these two sensors was calculated. All MEG values reported for this experiment are measurements of RMS amplitude and latency.

3. Results

3.1. Reaction times

The 2 (lexicality) × 2 (probability) × 2 (frequency) × 2 (density) ANOVAs were performed for reaction times and for accuracy. The Scheffe post hoc test was used to test for significant comparisons. An ANOVA on reaction times showed a significant overall effect of frequency ($F(1,15) = 127.06, p \leqslant .0001$). As Table 1 shows, high frequency words ($\bar{x} = 690.5$) were responded to more quickly than low frequency words ($\bar{x} = 739.9$). Instead of the expected main effect of density on RT, density was involved in a significant interaction with frequency ($F(1,15) = 5.677, p < .03$), as illustrated in Fig. 2. For high frequency words, high neighborhood density words elicited slightly delayed reaction times relative to low density words ($\bar{x} = 694.4$ and $\bar{x} = 686.6$, respectively), though the Scheffe's test revealed this effect to be unreliable ($p < .25$). For low frequency words, a significant effect occurred in the opposite direction (high density—$\bar{x} = 733.4$, low density—$\bar{x} = 746.5$ $p < .04$).

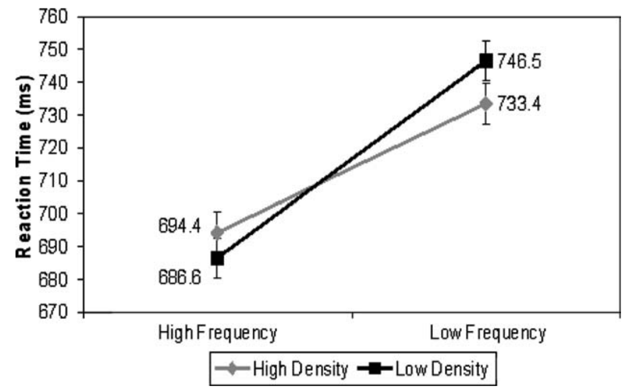

Fig. 2. Frequency × density interaction on reaction time.

Analyses of accuracy, revealed a significant effect of frequency on error rate. High frequency words were responded to incorrectly less often than low frequency words ($\bar{x} = 0.7$ vs. $\bar{x} = 2.1$) ($F(1,15) = 68.367$, $p \leqslant .0001$). No other significant effect of error rate was observed.

3.2. M350

3.2.1. Latency

The only significant effect on M350 latency was an effect of phonotactic probability. High probability items (words and non-words) elicited earlier M350 latencies

Table 1
The main effects of frequency, density, and probability on mean reaction times and error rates ($n = 16$)

Stimulus	Measure					
	RT			Accuracy		
	High	Low		High	Low	
	M (*SD*)	*M* (*SD*)	*F*	*M* (*SD*)	*M* (*SD*)	*F*
Frequency	690.5 (100.49)	739.9 (50.94)	127.06***	1.75 (0.85)	5.4 (1.76)	68.37***
Density	713.9 (130.31)	716.5 (123.10)	0.37	3.7 (1.46)	3.45 (1.57)	0.38
Probability	713.1 (121.19)	717.3 (131.99)	0.92	3.8 (1.49)	3.35 (1.54)	0.96

*** $p \leqslant .0001$.

Table 2
The main effects of frequency, density and probability on M350 and M250 latencies

Stimulus	Measure					
	M250 Latency			M350 Latency		
	High	Low		High	Low	
	M (*SD*)	*M* (*SD*)	*F*	*M* (*SD*)	*M* (*SD*)	*F*
Frequency	280.1 (21.73)	275.4 (23.14)	0.11	345.5 (22.48)	359.8 (23.14)	1.08
Density	274.8 (24.03)	280.7 (22.12)	0.16	355.1 (22.86)	350.2 (23.14)	1.79
Probability	283.2 (23.51)	272.3 (22.28)	9.21**	348.5 (23.20)	356.8 (23.04)	5.33*

* $p \leqslant .05$.
** $p \leqslant .01$.

than low probability items ($\bar{x} = 348.5$ vs. $\bar{x} = 356.8$) ($F(1, 15) = 5.331$, $p < .03$). Table 2 shows this effect.

3.2.2. Amplitude

There were no significant main effects on M350 amplitude. There was a significant interaction between probability and density ($F(1, 15) = 7.815$, $p < .008$), (see Fig. 3). For the low probability stimuli, high density items ($\bar{x} = 149.3$) elicited greater M350 amplitudes than low density items ($\bar{x} = 135.2$). However, for the high probability items, high density items ($\bar{x} = 128.6$) evoked smaller M350 amplitudes than low density stimuli ($\bar{x} = 139$). Scheffe's post hoc tests revealed two significant differences were reliable: high probability, high density items had smaller amplitudes than low probability, high density items ($p < .002$) and low probability, high density items had greater amplitudes than low probability, low density items ($p < .03$).

3.3. M250

M250 latencies were significantly affected by phonotactic probability (see Table 2). However, the effect was the opposite of that observed for M350 latencies. High probability items elicited significantly later M250 latencies than low probability items ($\bar{x} = 283.2$ vs. $\bar{x} = 272.3$) ($F(1, 15) = 9.207$, $p = .0044$). When a separate ANOVA was performed considering only the word stimuli, the probability effect on latency was not significant, but probability did have a significant effect on M250 amplitudes. High probability words ($\bar{x} = 142.3$) elicited higher amplitude M250s than low probability words ($\bar{x} = 126.6$) ($F(1, 15$, $p = .017$). An ANOVA including only non-word stimuli found no significant effects or interactions.

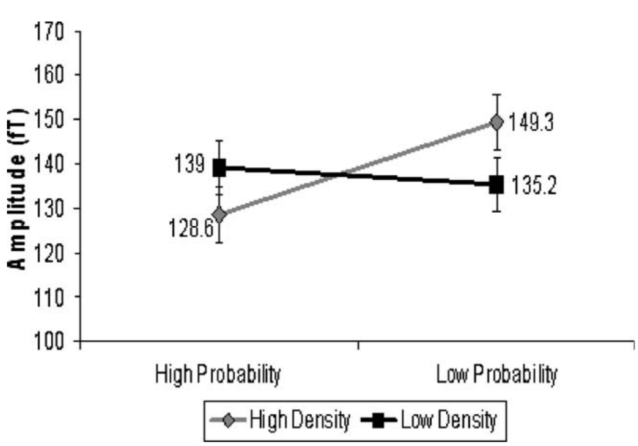

Fig. 3. Probability × density interaction on M350 amplitude.

3.4. M170

In contrast to the M350 and M250 none of the stimulus variables manipulated in this experiment had any significant effect on either the latencies or amplitudes of the M170 component.

4. Discussion

By independently manipulating stimulus properties that are usually highly correlated with each other, this experiment reveals evidence that these stimulus properties interact with each other in lexical processing in more complicated ways than has been previously observed. In particular, the results of this experiment suggest two areas requiring further investigation.

4.1. Frequency

Looking first at reaction time as a dependent measure, lexical frequency effects are partly what we expect them to be—high frequency items are responded to much more quickly than low frequency items. This is exactly the result that previous behavioral and MEG experiments predict. However, two complications emerge from the results of this study.

First, the interaction of density with frequency is an unexpected result for two reasons. One: There is no main effect of density on reaction time even when we consider only the nonword items. This is surprising considering the results of previous studies (Luce, Goldinger, Auer, & Vitevitch, 2000; Luce & Large, 2001; Luce & Pisoni, 1998; Vitevitch & Luce, 1999). However, the density measure used in this experiment is not frequency weighted. Only the absolute number of neighbors of each item was calculated, the relative frequency of those neighbors was not factored into the measure. This was necessary in order to vary the stimuli simultaneously by density, frequency and probability, but is not consistent with the above mentioned studies. Two: Density does seem to matter as expected when it comes to the low frequency words—high density items are responded to slower than low density words. But this inhibitory effect disappears with respect to the high frequency words. Apparently the processing advantage conferred on high frequency items is strong enough that post-activation competition is not sufficient to delay recognition.

Second, the highly significant main effect of frequency on reaction time makes the unreliability of such an effect on the M350 latency surprising. There is a nearly 50 ms difference in the mean latency of the M350 response to high vs. low frequency items, but this effect is only very marginally reliable ($F(1, 15) = 1.08, p < .3$). The large variance can be explained as a function of the peculiar design constraints of this experiment. Items had not only to vary in frequency, but also in probability and density, leading to items with a frequency of as low as 11 occurrences per million being classified as high frequency.

However, given that Embick et al. (2001) found a clear correlation between the facilitative effect of high frequency on reaction time and the facilitative effect on the latency of the M350 the results of the current study do raise the question as to whether there are any other factors at work responsible for the large processing advantage enjoyed by the high frequency items.

4.2. Probability

The observed effects of probability in this experiment also raise questions. The main effect of probability on the M350 confirms the conclusions of Pylkkänen et al. (2002) that sublexical frequency affects the timing of the M350. But the interaction of probability with density on the M350 amplitude calls into question the other conclusion of that paper, i.e., that density does not affect the M350 and that therefore that component indexes initial, pre-competition stages of lexical access. The result observed in the current experiment is that stimuli with high neighborhood densities and low bi-phone probabilities show inhibited processing at a stage in processing which Pylkkänen et al. (2002) show is not sensitive to competition induced inhibition. However, as mentioned above in the discussion on frequency, the density measure used in this experiment was different from the frequency weighted measure of density used by Pylkkänen and colleagues.

Another interesting result of the current experiment is the effect of probability on the M250. Both for amplitude (across all items) and latency (for words only), high probability stimuli actually inhibited the M250 response. Not only is this result in direct conflict with the facilitation found by Pylkkänen et al. (2002), it is also surprising given that high probability items facilitated processing at the stage indexed by the M350 in the current experiment. The inhibition found here suggests that there is an earlier effect of competition that previously supposed.

References

Adachi, Y., Shimogawara, M., Higuchi, M., Haruta, Y., & Ochiai, M. (2001). Reduction of nonperiodical extramural magnetic noise in MEG measurement by continuously adjusted least squares method. *IEEE Transactions on Applied Superconductivity*, ASC2000, Virginia Beach.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 252*, 257–271.

Embick, D., Hackl, M., Schaeffer, J., Kelepir, M., & Marantz, A. (2001). A magnetoencephalographic component whose latency reflects lexical frequency. *Cognitive Brain Research, 10*(3), 345–348.

Helenius, P., Salmelin, R., Service, E., & Connolly, J. F. (1998). Distinct time courses of word and context comprehension in the left temporal cortex. *Brain, 121*, 1133–1142.

Helenius, P., Salmelin, R., Service, E., & Connolly, J. F. (1999). Semantic cortical activation in dyslexic readers. *Journal of Cognitive Neuroscience, 11*(5), 535–550.

Koyama, S., Kakigi, R., Hoshiyama, M., & Kitamura, Y. (1998). Reading of Japanese Kanji (morphograms) and Kana (syllabograms): A magnetoencephalographic study. *Neuropsychologia, 36*(1), 83–98.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Kuriki, S., Takauchi, F., Fujimaki, N., & Kobayashi, T. (1996). Magnetoencephalographic study on the cerebral neural activities related to the processing of visually presented characters. *Cognitive Brain Research, 4*, 185–199.

Luce, P. A., Goldinger, S. D., Auer, E. T., & Vitevitch, M. S. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception and Psychophysics, 62*, 615–625.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19*, 1–36.

Luce, A., & Large, N. (2001). Phonotactics, density, and entropy in spoken word recognition. *Language and Cognitive Processes, 16*(5/6), 565–581.

Pylkkänen, L., & Marantz, A. (2003). Tracking the time course of word recognition with MEG. *Trends in Cognitive Sciences, 7*, 187–189.

Pylkkänen, L., Stringfellow, A., Flagg, E., & Marantz, A., (2000). A neural response sensitive to repetition and phonotactic probability: MEG investigations of lexical access. *Proceedings of Biomag 2000*, Helsinki, Finland.

Pylkkänen, L., Stringfellow, A., & Marantz, A. (2002). Neuromagnetic evidence for the timing of lexical activation: An MEG component sensitive to phonotactic probability but not to neighborhood density. *Brain and Language, 81*, 666–678.

Sekiguchi, T., Koyama, S., & Kakigi, R. (2000). The effect of word repetition on evoked magnetic responses in the human brain. *Japanese Psychological Research, 42*, 3–14.

Tarkiainen, A., Helenius, P., Hansen, P. C., Cornelissen, P. L., & Salmelin, R. (1999). Dynamics of letter string perception in the human occiptotemporal cortex. *Brain, 122*, 2119–2132.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40*, 374–408.