RESEARCH ON SPOKEN LANGUAGE PROCESSING
Progress Report No. 21 (1996-1997)
Indiana University

Some Observations on Neighborhood Statistics of Spoken English Words[1]

Shigeaki Amano[2]

Speech Research Laboratory
Department of Psychology
Indiana University
Bloomington, Indiana 47405

[1] This work is supported in part by NIH-NIDCD Research Grant DC00111 to Indiana University. I am grateful to Prof. David Pisoni and Gina Torretta who allowed me to use their database.
[2]NTT Basic Research Laboratories, 3-1 Morinosato Wakamiya, Atsugi, Kanagawa, 243-01 Japan.

Some Observations on Neighborhood Statistics of Spoken English Words

Abstract. Lexical neighborhoods were analyzed as a function of word length and phoneme position in terms of the neighborhood density, the mean neighborhood frequency, and the maximum neighborhood frequency. It was found that the neighborhood is very sparse in words more than six phoneme long, and that there are phoneme positional differences in the neighborhood, suggesting that the conventional lexical neighborhood may not be a good word candidate set for spoken word recognition.

Introduction

Research on spoken word recognition suggests that a word is processed interactively with other word candidates in a mental lexicon (e.g., McQueen, Norris, & Cutler 1994; Norris, McQueen, & Cutler 1995; Zwitserlood, 1989). Some recent models of spoken word recognition explicitly take account of the competition among word candidates. For example, the TRACE model (McClelland & Elman, 1986) and the SHORTLIST model (Norris, 1994; Norris, McQueen, & Cutler, 1995) incorporate lexical competition by adopting inhibitory connections among word candidates. On the other hand, the original cohort model (Marslen-Wilson & Welsh, 1978), the new cohort model (Marslen-Wilson, 1987), and the Neighborhood Activation Model (Luce, 1986) implicitly take account of the interactions among word candidates.

One of the findings used to support an interactive view is the effect of lexical neighborhood on word recognition. A neighborhood is a collection of words which have single phoneme substitution with a target word (e.g., Frauenfelder, 1990; Frauenfelder, Baayen, Hellwig, & Schreuder, 1993; Pisoni, Nusbaum, Luce & Slowiaczek, 1985).

Precisely speaking, there is another definition for the neighborhood. This definition includes words with single phoneme deletion or addition to the target words in addition to the substitution (e.g., Goldinger, 1989; Luce, 1986; Sommers, 1996). However, the former is used as the definition of neighborhood in this paper, because it is simpler than the latter and no substantial differences were found in the statistical characteristics of the two definitions (Frauenfelder, Baayen, Hellwig, & Schreuder, 1993).

The characteristics of the neighborhood have been described by three variables in previous studies. The first variable is the "density" which is the number of words in the neighborhood (e.g., Frauenfelder, Baayen, Hellwig, & Schreuder, 1993; Luce, 1986). The second variable is the "mean frequency" which is the averaged frequency of words in the neighborhood (e.g., Frauenfelder, Baayen, Hellwig, & Schreuder, 1993; Luce, 1986). The third variable is the "maximum frequency" which is the highest frequency of a word in the neighborhood. (Bard, 1990; Bard & Shillcock, 1993).

Several studies have shown that the neighborhoods have significant effects on word recognition. For example, neighborhood density and frequency negatively correlate with recognition rate of a target word (Luce, 1986). Neighborhood density has inhibitory effects on reaction times of lexical decision and naming for a target word (Goldinger, 1989). Low frequency words in neighborhood have negative priming effects on target word recognition (Goldinger, Luce, & Pisoni, 1989). Recognition of two-syllable target words (spondees) is affected by neighborhood characteristics of each syllable in the spondee (Cluff & Luce, 1990). And, older adults have difficulty recognizing a target word if density and frequency is high in neighborhood (Sommers, 1996).

One of the problems in these studies is that almost all of them used only monosyllabic words. The exception was Cluff and Luce (1990) who used two-syllable spondee words. However, they calculated the neighborhood for each syllable not for a entire word. Therefore, effects of neighborhood have not been investigated in multisyllabic words. The reason is probably that the neighborhood is not effective in the multisyllabic words. In the initial analysis in this study, this idea will be confirmed by examining the statistics of neighborhoods.

Another problem with the previous studies is concerned with an assumption concerning neighborhood characteristics. In the past, researchers have assumed that neighborhoods at each phoneme position within a word have equivalent effects on word recognition. That is, the conventional definition of a neighborhood does not take into account positional factors. Frauenfelder, Baayen, Hellwig, and Schreuder (1993) have already mentioned this point, but did not conduct any analyses to investigate these effects. If speech is processed in a left-to-right manner (Cole & Jakimik, 1980) and if a spoken word is recognized by reference to the mental lexicon in a left-to-right manner as some word recognition models claim (e.g., Marslen-Wilson & Welsh, 1978; McClelland & Elman, 1986), then each phoneme position may contribute to word recognition differently. For instance, the first position might have a larger contribution to word recognition than other phoneme positions in a word.

Amano, Torretta, and Luce (1997) have found positional neighborhood effects. They conducted correlation analyses between word identification rate and three neighborhood variables; the density, the mean log frequency, and the maximum log frequency at each phoneme position in CVC English words. They reported that the first phoneme position had greater negative effects on word recognition than any of the other phoneme positions.

However, Amano, Torretta, and Luce (1997) used CVC words in their study. As a consequence, the length of words was restricted to only three phonemes. It is unknown whether their results are applicable to other word lengths. A second analysis focused on this point in order to provide indirect evidence that the first phoneme position is also effective in other word lengths. Thus, the statistics of the positional neighborhood was examined in longer words.

Analysis 1

Method

A computerized dictionary (Nusbaum, Pisoni, & Davis, 1984) was used for all the following analyses. The dictionary contains 19,295 words with Kucera and Francis (1967) word count. However, only 19,152 words with two to twelve phoneme long were used for the following analyses, because there is no neighborhood for words with more than 13 phoneme long. The analyses were conducted as a function of word length. Independent variables were the number of words, the log frequency of a target word, the neighborhood density, the mean neighborhood log frequency, and the maximum neighborhood log frequency.

Results and Discussion

Figure 1 shows the distribution of the number of words as a function of word length. It shows that the words which are five phoneme long are the most frequent. The two phoneme long words and words in more than 10 phoneme long are not very frequent.

Insert Figure 1 about here.

Figure 2 shows the distribution of log frequencies of a target word. The results indicate that two phoneme words are the most frequent and that word frequency exponentially decreases as word length increases.

Insert Figure 2 about here.

Figure 3 shows the distribution of neighborhood density as a function of word length. The neighborhood is very dense for two and three phoneme words, but becomes very sparse in words with more than six phonemes. Note that monosyllabic words are up to seven phonemes long in English and that such words are frequent in three and four phoneme long. These results suggest that neighborhoods are present for monosyllabic words but are virtually non-existent for multisyllabic words. Therefore, the neighborhood is only effective for monosyllabic words but not for multisyllabic words.

Insert Figure 3 about here.

This conclusion is supported by the distribution of the mean neighborhood log frequency (Figure 4) and the maximum neighborhood log frequency (Figure 5). These two variables are almost zero for words with more than six phonemes. The findings suggest that neighborhood frequency is not effective for those words. Therefore, neighborhood is not effective either in terms of density or frequency for multisyllabic words.

However, as shown by Cluff and Luce (1990), if neighborhood is calculated at each syllable in multisyllabic words, it affects recognition of multisyllabic words. Therefore, the effects of neighborhood are not straightforward for multisyllabic words. It is effective via a syllable but not via an entire word.

Insert Figure 4 about here.

Insert Figure 5 about here.

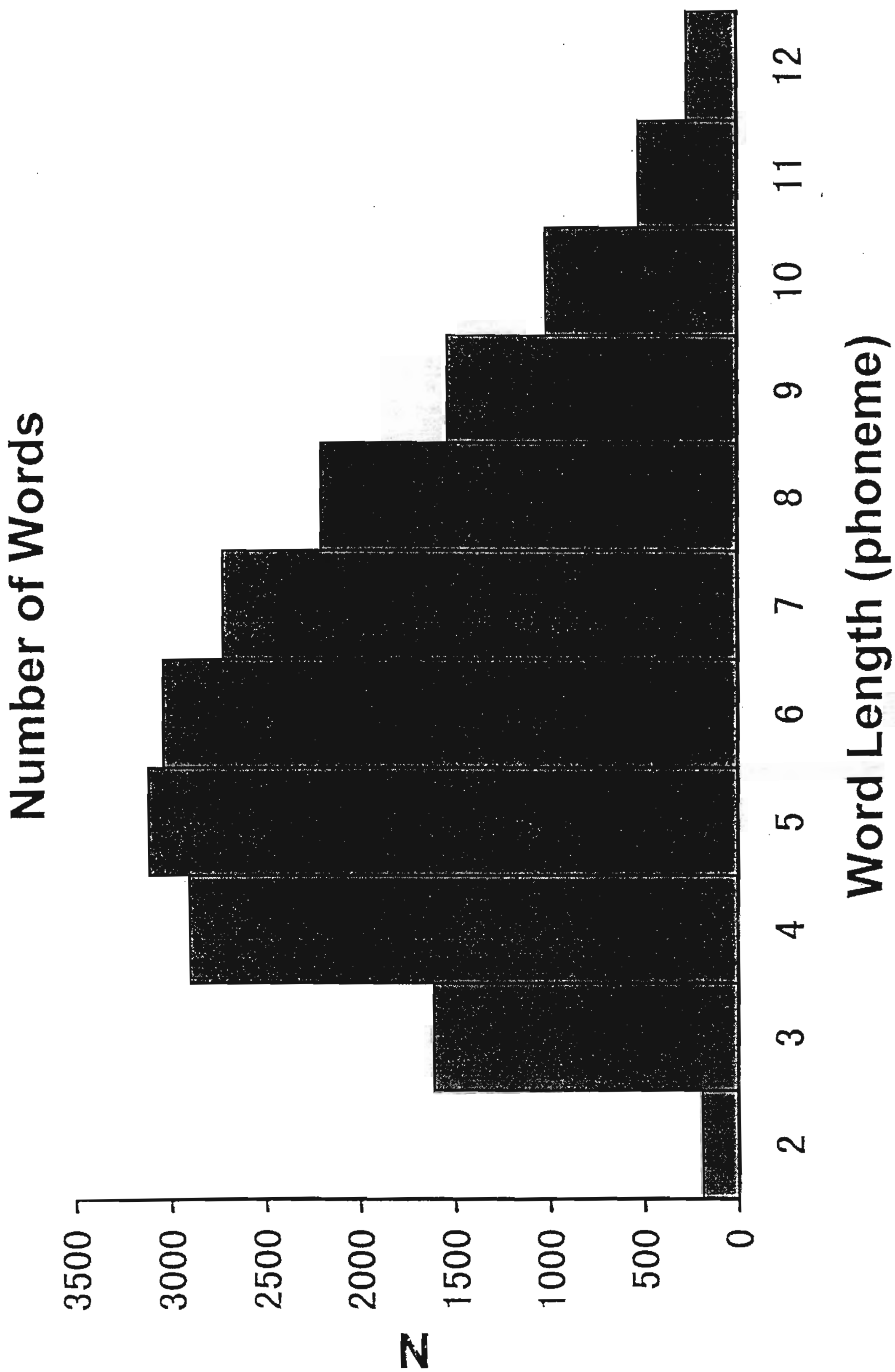

Figure 1. The number of words as a function of word length.

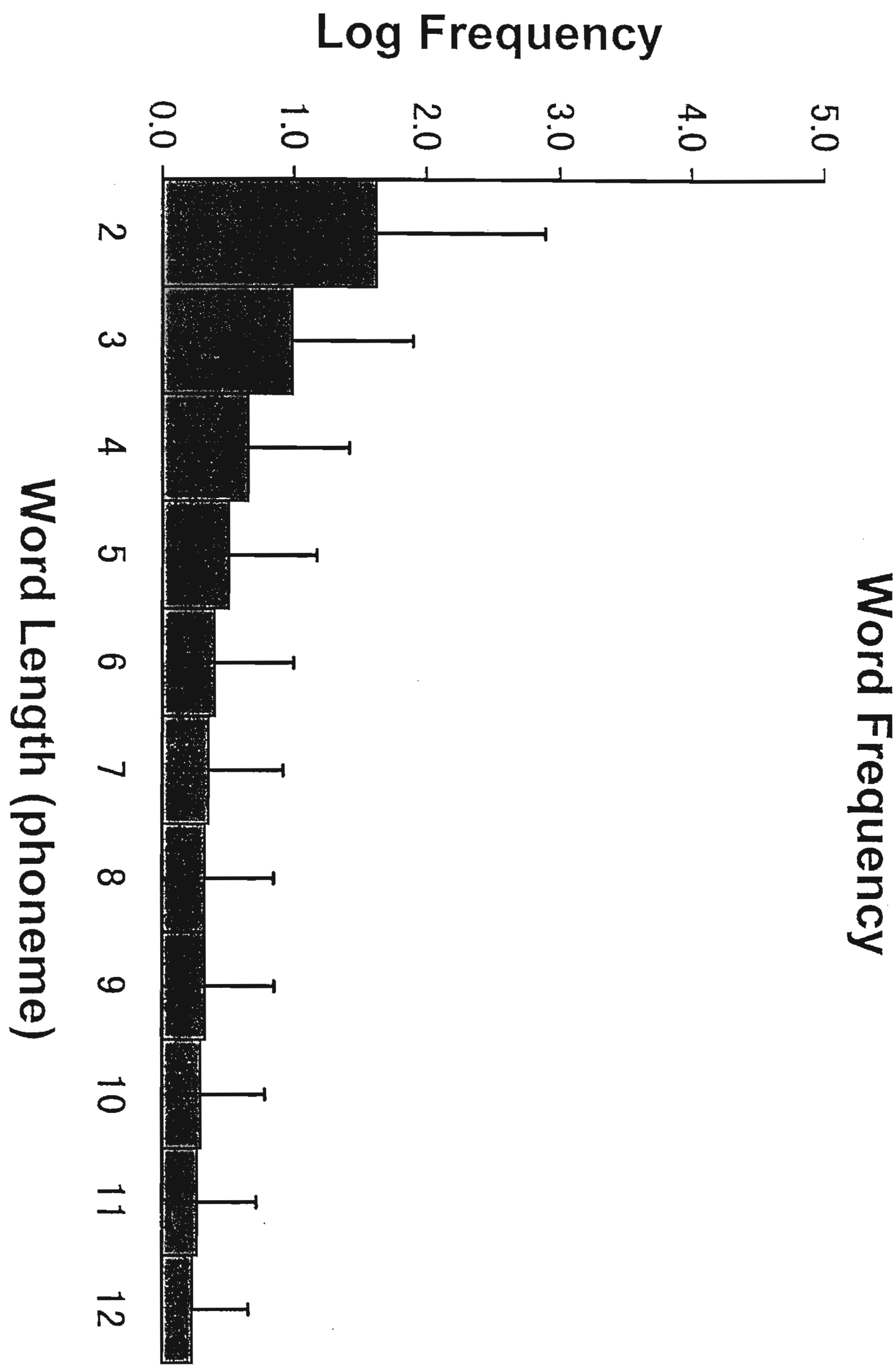

Figure 2. Word frequency of a target word as a function of word length.

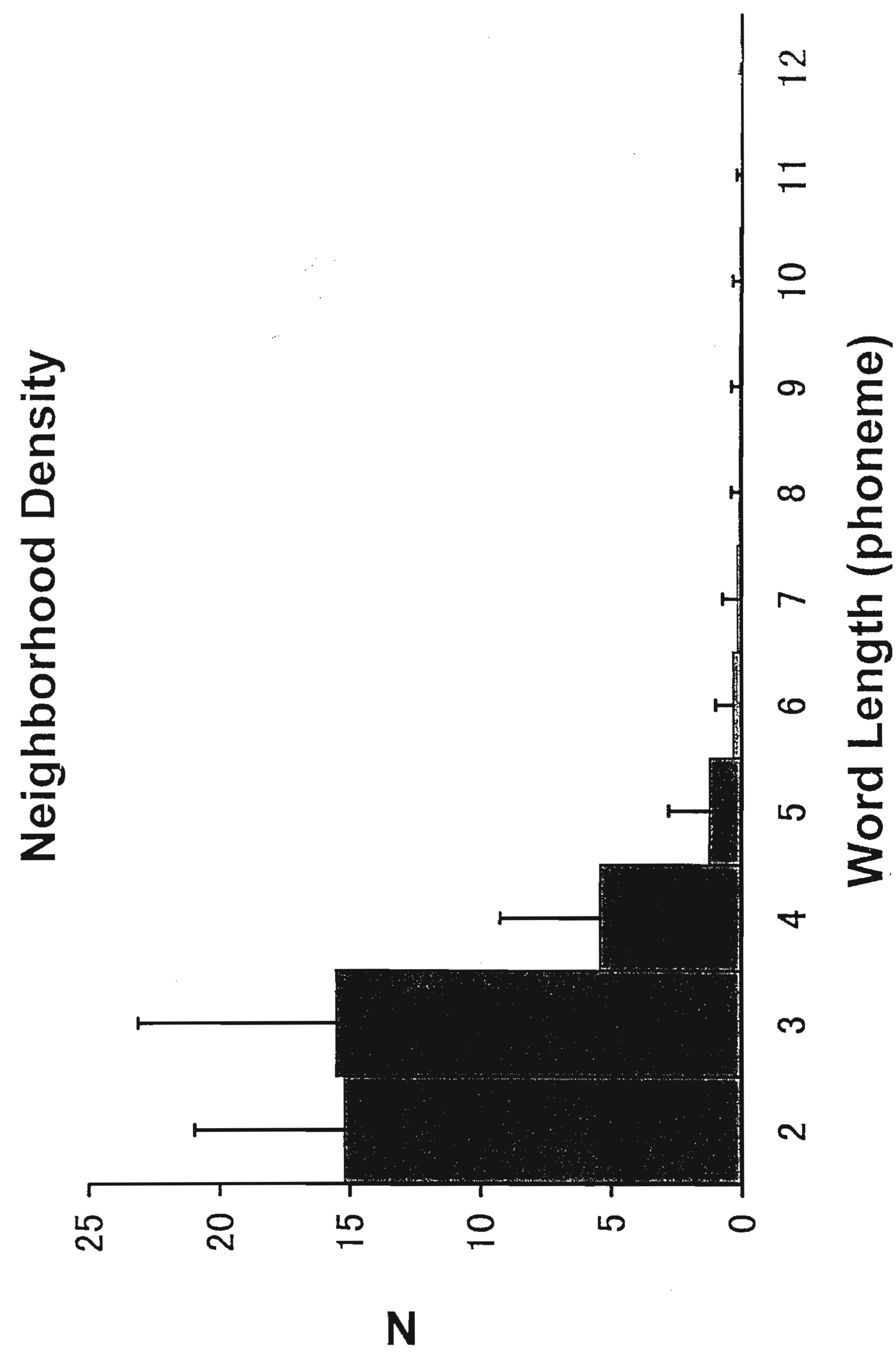

Figure 3. Neighborhood density as a function of word length.

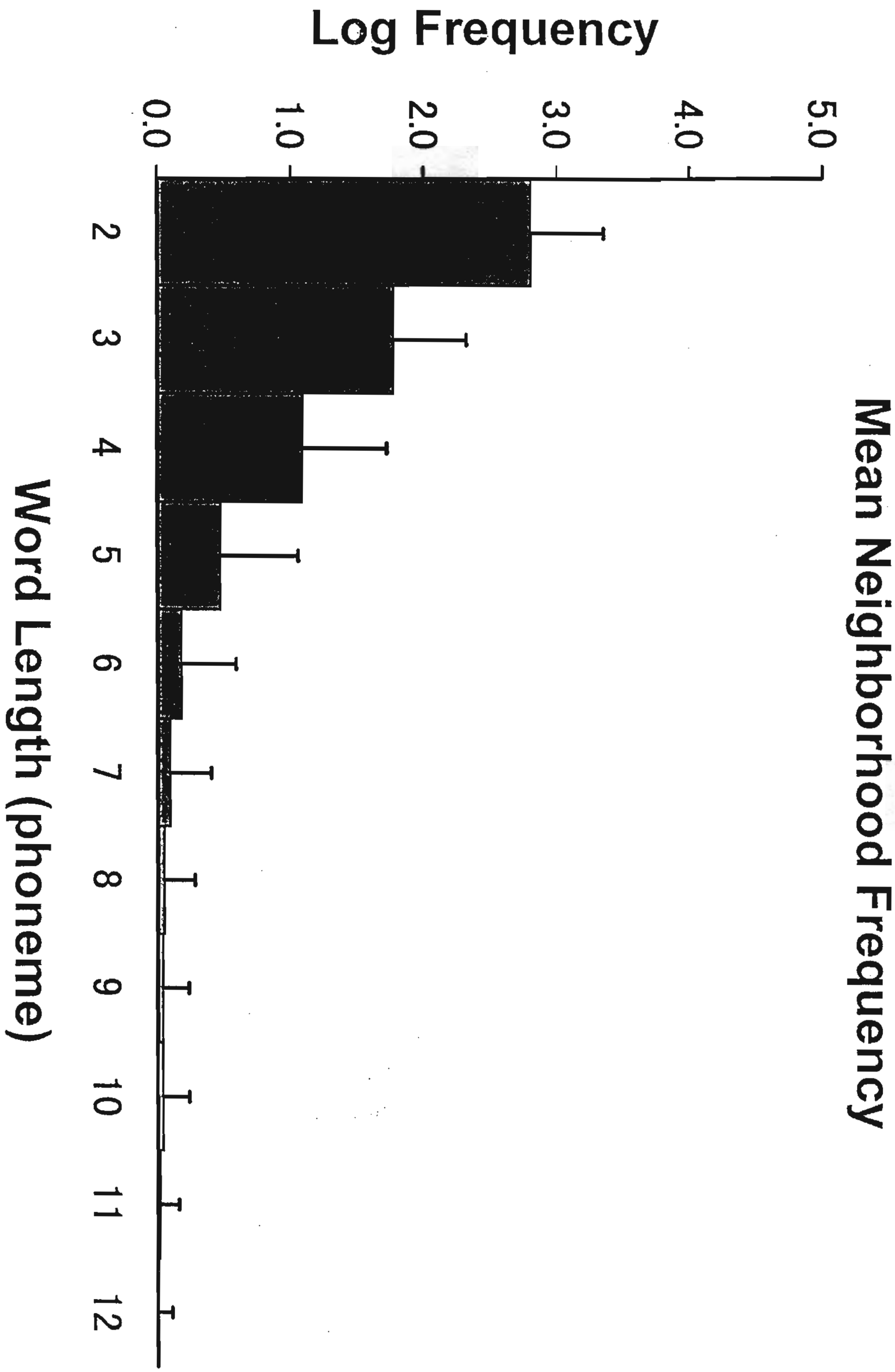

Figure 4. Mean neighborhood log frequency as a function of word length.

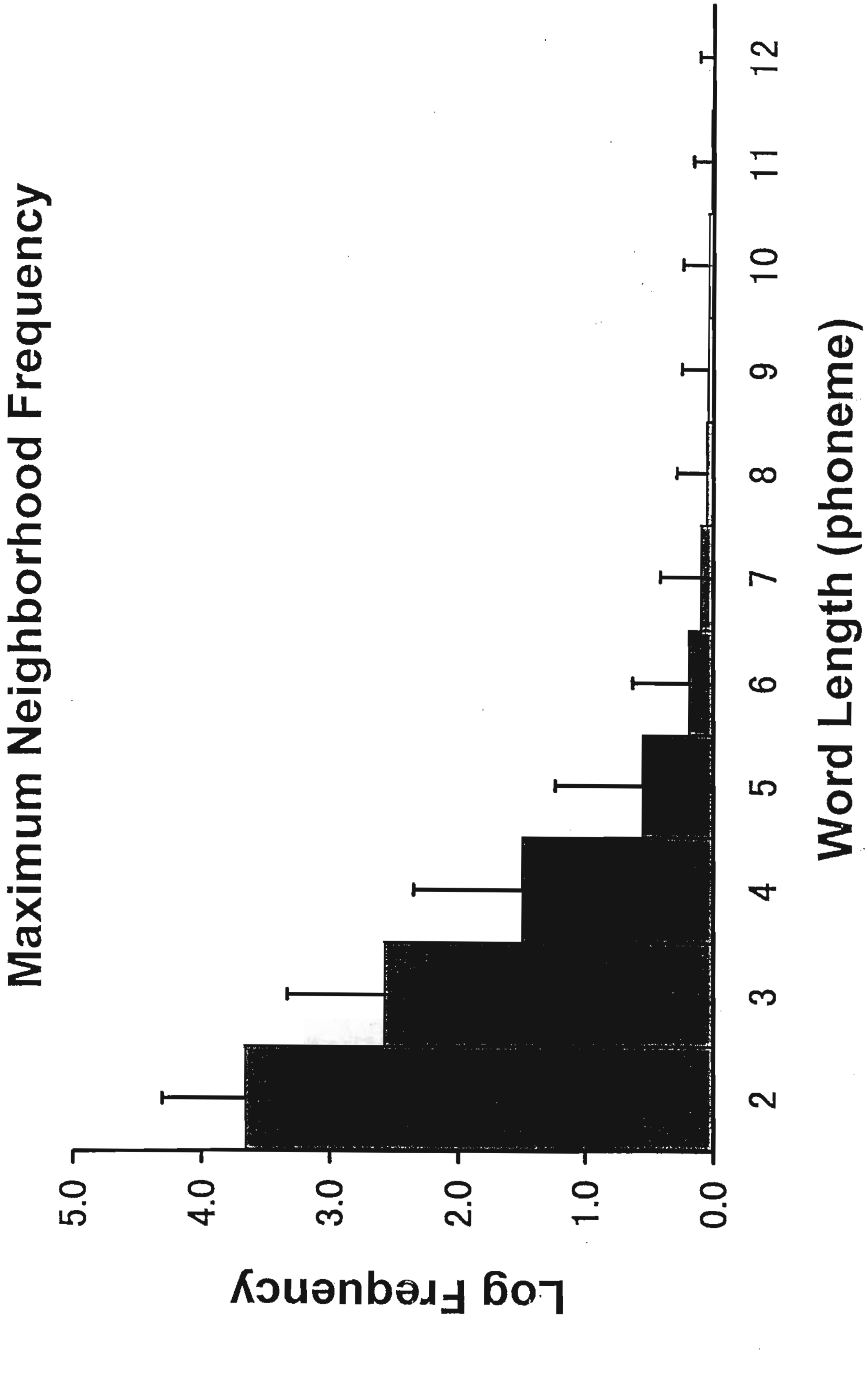

Figure 5. Maximum neighborhood log frequency as a function of word length.

Analysis 2

Method

The same word set as Analysis 1 was used for another series of analyses. This analysis was concerned with statistics of positional neighborhoods (Amano, Torretta, & Luce, 1997) which have single phoneme substitution with a target word at a particular phoneme position. For example, if a target word is "cat," then "bat," "fat," "hat," and "pat" are in the positional neighborhood at the first phoneme position. The density, the mean log frequency, and the maximum log frequency of the positional neighborhood were calculated for every phoneme position in two to twelve phoneme long words.

Results and Discussion

Figures 6, 7, and 8 show the distribution of the density, the mean log frequency, and the maximum log frequency of the positional neighborhood. Like the conventional neighborhood which was examined in Analysis 1, the density of the positional neighborhood is very low for words with more than six phonemes (Figure 6). The mean log frequency (Figure 7) and the maximum log frequency (Figure 8) of positional neighborhood are also very low for these words. Therefore, it is suggested that the positional neighborhood is not effective for words with more than six phonemes. This finding suggests that the positional neighborhood is not as effective in multisyllabic words as the conventional neighborhood as shown in Analysis 1.

Amano, Torretta, and Luce (1997) have shown that the positional neighborhood at the first phoneme position is more effective as a predictor for lexical competition than that at other phoneme positions. They used three phoneme words with CVC pattern.

The present analyses showed that there was a significant difference among the phoneme positions in three phoneme long words in terms of the neighborhood density, $F(2,4896) = 242.69$, $p < .0001$, the mean neighborhood log frequency, $F(2,4896) = 19.54$, $p < .0001$, and the maximum neighborhood log frequency, $F(2,4896) = 46.21$, $p < .0001$. The HSD test between a pair of positions showed that the first phoneme position is greater than the other phoneme positions in almost all cases in terms of all neighborhood variables with 5% significance level. The only one exception was that there was no significant difference between the first phoneme position and the third phoneme position in term of the mean neighborhood log frequency.

The results of Amano, Torretta, and Luce (1997)'s study might be due to the superiority of the first phoneme position in the positional neighborhood. In that case, the effect of positional neighborhood at the first phoneme position on spoken word recognition may be observed not only in three phoneme words but also in words of other lengths.

Insert Figure 6 about here.

Insert Figure 7 about here.

Insert Figure 8 about here.

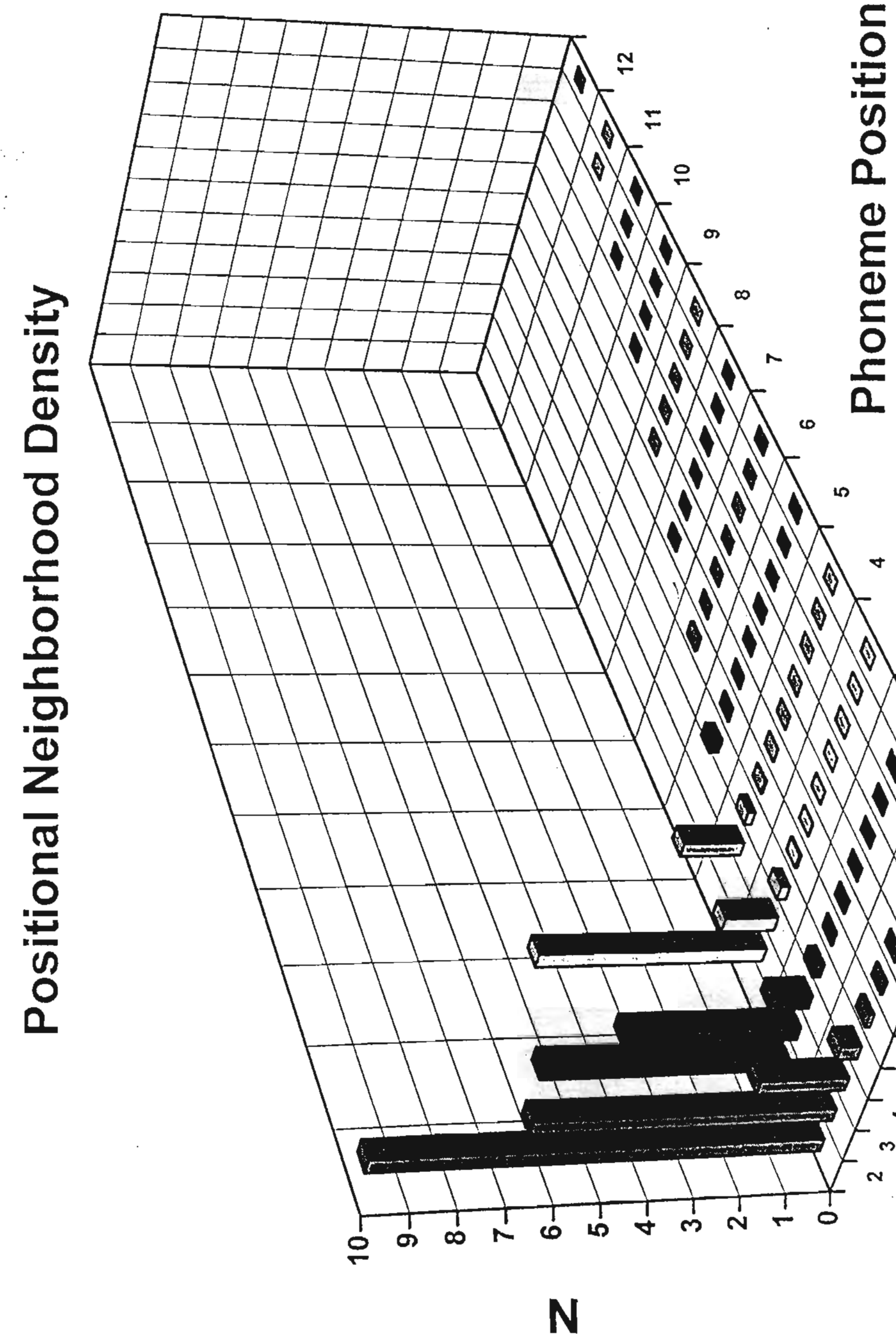

Figure 6. Positional neighborhood density.

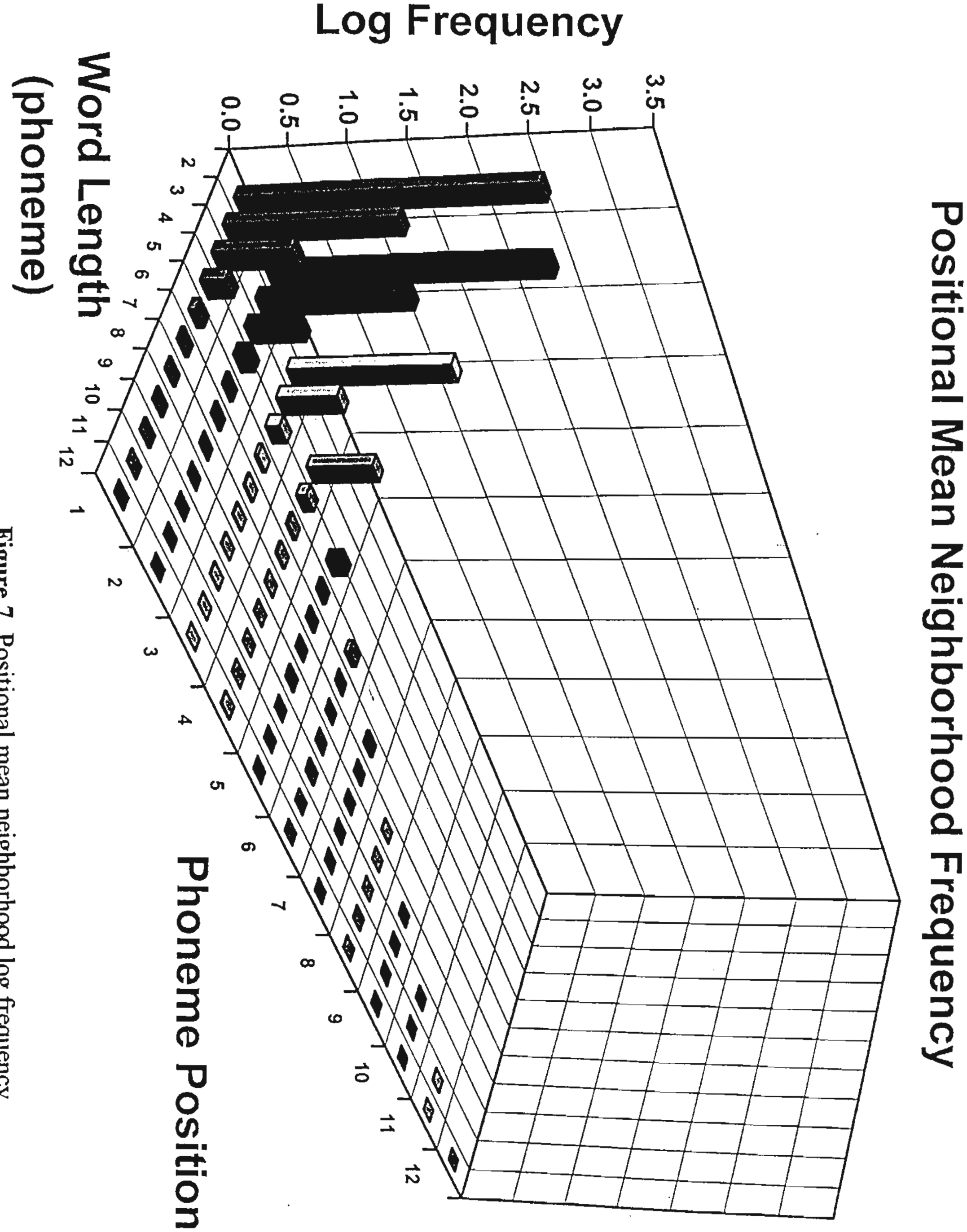

Figure 7. Positional mean neighborhood log frequency.

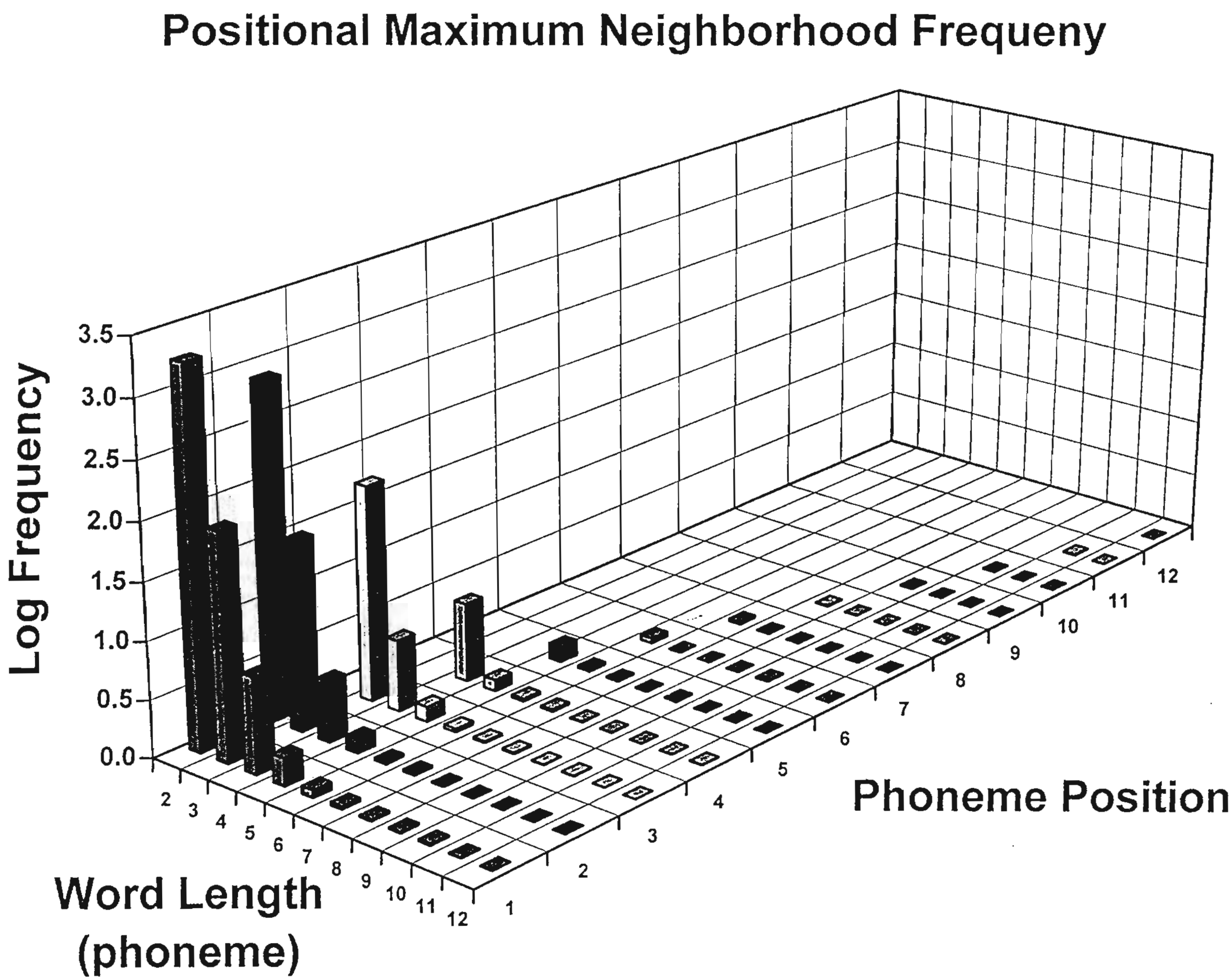

Figure 8. Positional maximum neighborhood log frequency.

Support for this proposal was obtained as follows. There was a significant main effect of phoneme position in almost all lengths with 1% significance level (exceptions were 12 phoneme words for the neighborhood density, 2, 11, and 12 phoneme words for the mean neighborhood log frequency, and 11, and 12 phoneme words for the maximum neighborhood log frequency). In addition, the first phoneme position is significantly greater than all other phoneme positions in two to seven phoneme words in terms of the neighborhood density, in three to six phoneme words in terms of the mean neighborhood log frequency, and in two to six phoneme words in terms of the maximum neighborhood log frequency. This means that the first position has greater value than other phoneme positions in the word length where the positional neighborhood is virtually effective. Therefore, it is suggested that the first phoneme position contributes more to spoken word recognition than other phoneme positions not only in three phoneme words (Amano, Torretta, & Luce, 1997) but also in words of other lengths.

Conclusion

Analysis 1 suggested that the neighborhood is not effective in words with more than six phonemes, because the neighborhood is very sparse in those words. Analysis 2 showed that there are phoneme positional differences in the neighborhood and that the first phoneme position has greater value than other phoneme positions in terms of the neighborhood density, the mean neighborhood frequency, and the maximum neighborhood frequency.

These findings indicate that the conventional neighborhood has some difficulties in explaining lexical competition in spoken English words. It cannot explain the competition between long words, and it does not take into account positional differences. Therefore, the neighborhood is not suitable for the word candidate set for spoken word recognition. Further research is required to find a better word candidate set.

References

Amano, S., Torretta, G. M., & Luce, P. A. (1997). Positional neighborhood effects on spoken word recognition. *The Journal of the Acoustical Society of America*, **101**, Part 2, 4aSC9, 3155.

Bard, E. G. (1990). Competition, lateral inhibition, and frequency: Comments on the chapters of Frauenfelder and Peeters, Marslen-Wilson, and others. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 185-210). Cambridge: MIT Press.

Bard, E. G., & Shillcock, R. C. (1993). Competitor effects during lexical access: Chasing Zipf's tail. In G. T. M. Altmann, & R. Shillcock (Eds.), *Cognitive models of speech processing: The second Sperlonga meeting* (pp. 235-275). Hillsdale: LEA.

Cluff, M. S., & Luce, P. A. (1990). Similarity neighborhoods of spoken two-syllable words: Retroactive effects on multiple activation. *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 551-563.

Cole, R. A., & Jakimik, J. (1980). A model of speech perception. In R. A. Cole (Ed.), *Perception and production of fluent speech* (pp. 133-164). Hillsdale: LEA.

Frauenfelder, U. H. (1990). Structure and computation in the mental lexicon. In H. Haken & M. Stadler (Eds.), *Synagetics of cognition* (pp. 406-414). Berlin: Springer-Verlag.

Frauenfelder, U. H., Baayen, R. H., Hellwig, F. M., & Schreuder, R. (1993). Neighborhood density and frequency across languages and modalities. *Journal of Memory and Language*, **32**, 781-804.

Goldinger, S. D. (1989). Neighborhood density effects for high frequency words: Evidence for activation-based models of word recognition *Research on Speech Perception, Progress Report 15*, pp. 163-186. Bloomington, IN: Indiana University, Speech Research Laboratory.

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language*, **28**, 501-518.

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

Luce, P. A. (1986). Neighborhoods of words in the mental lexicon. *Research on Speech Perception, Technical Report* **6**. Bloomington IN: Indiana University, Speech Research Laboratory.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, **25**, 71-102.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, **10**, 29-63.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, **18**, 1-86.

McQueen, J. M., Norris, D., & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **20**, 621-638.

Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition*, **52**, 189-234.

Norris, D., McQueen, J. M., & Cutler, A. (1995). Competition and segmentation in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 1209-1228.

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words. *Research on Speech Perception, Progress Report 10*, pp. 357-376. Bloomington, IN: Indiana University, Speech Research Laboratory.

Pisoni, D. B., Nusbaum, H. C., Luce, P. A., & Slowiaczek, L. M. (1985). Speech perception, word recognition and the structure of the lexicon. *Speech Communication*, **4**, 75-95.

Sommers, M. S. (1996). The structural organization of the mental lexicon and its contribution to age-related declines in spoken-word recognition. *Psychology and Aging*, **11**, 333-341.

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition*, **32**, 25-64.