

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

Lexical Competition in Spoken English Words¹

Shigeaki Amano²

*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 David Pisoni and Gina Torretta who allowed me to use their database. They also gave me valuable comments and suggestions on an earlier draft of this paper.

² NTT Basic Research Laboratories, 3-1 Morinosato-Wakamiya, Atsugi, Kanagawa, 243-01 Japan. E-mail: amano@av-hp.brl.ntt.co.jp

Lexical Competition in Spoken English Words

Abstract. The “rime cognate,” which contains the same phoneme sequence as the rime of a given word, is proposed as a new lexical competitor set for spoken word recognition. Partial correlation analyses showed that the rime cognate has a greater negative contribution to the identification of English spoken words than the lexical neighborhood and the word-initial cohort which had been commonly used as a lexical competitor set. The analyses also showed that the rime cognate includes the reliable part of the neighborhood and the cohort, which indicates that the important parts of these two different competitor sets are integrated into the rime cognate.

Introduction

Research on spoken word recognition suggests that multiple word candidates are activated during the recognition process, and that these candidates compete with each other. Some empirical support for activation of multiple word candidates has been obtained in a cross-modal semantic priming task. For instance, Shillcock (1990) has shown that, when a carrier word is processed, an embedded word in the carrier word is also activated. For example, “bone” is activated when “trombone” is processed. Connine, Blasko, and Wang (1994) also used the cross-modal semantic priming task to investigate spoken word recognition. They found that multiple word candidates are activated when a spoken word has an ambiguous initial phoneme. Zwitserlood (1989) has also shown that sentential-semantic contexts are used for selecting one of the activated word candidates.

Although these studies indicate that the multiple word candidates are activated during the recognition process, there is further support that the multiple word candidates compete with each other. For example, McQueen, Norris, and Cutler (1994) measured reaction times for detecting the monosyllabic target word which is embedded in the second syllable in a carrier nonword with a weak-strong syllable pattern. They found that the reaction times for the target word are longer when the carrier nonword corresponded to the beginning of an other multisyllabic word than when it did not. For example, detection of “mess” in “domes” (which is a part of “domestic”) required longer reaction time than detection of “mess” in “nemess,” indicating that “domestic” competes with “mess.” Their results suggest that the multisyllabic word which shares its beginning part with the carrier nonword competes the monosyllabic target word.

Evidence for such lexical competition has also been obtained in the studies of phonological priming (Hamburger & Slowiaczek, 1996, Slowiaczek & Hamburger, 1992). Slowiaczek and Hamburger (1992) have found interference effects on reaction time for shadowing of the target word when the prime has an initial three phoneme overlap with the target word, suggesting the presence of competition between the target word and the prime.

Other evidence of lexical competition has been reported by Vroomen and de Gelder (1995). They found that facilitative cross-modal priming effects are larger for auditory primes with few or no competitors than for auditory primes with many competitors. Their results suggest that competitors have an inhibitory effect on word recognition according to their numbers. They also found that the difference of priming effects according to the number of competitors disappeared when there was no interstimulus

interval between auditory prime and visual target word. This result suggests that lexical competition requires some amount of time to be effective.

Several recent models of spoken word recognition explicitly assume competition among word candidates. For example, the TRACE model (McClelland & Elman, 1986) and the SHORTLIST model (Norris, 1994; Norris, McQueen, & Cutler, 1995) both incorporate lexical competition by adopting inhibitory connections among word candidates. On the other hand, other models such as the original cohort model (Marslen-Wilson & Welsh, 1978), the new cohort model (Marslen-Wilson, 1987), and the Neighborhood Activation Model (Luce, 1986) do not have such inhibitory relationship among word candidates. However, as pointed by Frauenfelder (1996), these models incorporate the competition implicitly, because they use a decision rule which determines the "winning" candidates by taking account of the status of all other word candidates.

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

Although these variables has been used commonly, there are some contradictions about which variable has the largest contribution to word recognition (e.g., Bard, 1990; Bard & Shillcock, 1993). More importantly, there has been little general agreement on the set of word candidates. That is, there is a discrepancy on what kind of words are included in the set. Previous studies assume at least two different candidate sets for spoken word recognition; the lexical neighborhood and the word-initial cohort.

A lexical neighborhood is defined as 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). Other definition for the neighborhood 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). The former definition is used as the neighborhood in this paper, because it is simpler than the latter, and no substantial differences were found in the statistical characteristics between the two definitions (Frauenfelder, Baayen, Hellwig, & Schreuder, 1993).

Several recent studies indicate that the neighborhoods have significant effects on word recognition performance. 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 in lexical decision and naming for a target word (Goldinger, 1989). Low frequency words in a 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 word (Cluff & Luce, 1990). Older adults have difficulty recognizing a target word if density and frequency is high in neighborhood (Sommers, 1996). Taken together, all these studies clearly indicate that the lexical neighborhood affects its recognition.

In contrast to the role of the neighborhood, several studies have shown that reliable word candidate set is the cohort which is a set of words sharing the initial part of phoneme sequences with a target word (e.g., Marslen-Wilson & Welsh, 1978; Marslen-Wilson, 1989; Bard & Shillcock, 1993). For example, Marslen-Wilson (1990) found that the isolation point (IP) of a word in the gating task (Grosjean, 1980) is

later for a word with high frequency competitors in its cohort than it is for a word with low frequency competitors. His results show that word candidates in the cohort affect the word recognition and that the amount of effect is a function of their frequency.

In the definition of the neighborhood and the cohort, an English word is treated as a simple sequence of phonemes. However, some linguistic studies argue that words can be sub-divided into syllables. The syllable, in turn, has two main parts, an onset and a rime (Cairns & Feinstein, 1982; Fudge, 1969; Halle & Vergnaud, 1980). The onset contains the initial consonant cluster of the syllable. The rime contains the peak (vowel nucleus) and the coda (final consonant cluster).

Evidence for the division of the syllable into onset and rime components has been obtained in numerous psycholinguistic studies. For example, support has been found in the studies of speech errors (MacKay, 1972; Stemberger & Treiman, 1986), short-term memory of syllables (Treiman & Danis, 1988), word games for syllable blending (Treiman, 1983, 1986), and for syllable dividing and transforming (Treiman, 1986). Further support has been obtained from children's performance of word games for syllable transforming (Treiman, 1985), rime identification (Lenel & Cantor, 1981; Treiman & Zukowski, 1996), and word categorization (Kirtley, Bryant, MacLean, & Bradley, 1989).

Given the psycholinguistic importance of the onset and the rime, two alternatives can be considered as word candidates for lexical competition. One is the set of words which share the onset, the other is the set of words which are related to each other through the rime.

Broadly speaking, the onset-sharing set corresponds to the word-initial cohort. When a word begins with a consonant, the cohort which shares the initial part of word coincides with the onset-sharing set. When a word begins with a vowel, the cohort does not coincide with the onset-sharing set. However, words beginning with a consonant are more frequent than the words beginning with a vowel in English. 79.7% of words begin with a consonant when a search is conducted using a computerized dictionary (Nusbaum, Pisoni, & Davis, 1984) which contains 19,295 words. This fact means that the cohort roughly corresponds to the onset-sharing set. Therefore, it is not necessary to consider it here.

On the other hand, the set of words which are related to each other through the rime has not been investigated. This set is hereafter called a "rime cognate." The precise definition of the rime cognate is a set of the words which contain the same sequence of phonemes with the rime of a given word. The lexical matching procedure used to produce the rime cognate is hereafter called rime matching strategy (RMS).

For example, if a word "cat" is given, its rime cognate produced by the RMS contains, for instance, "hat", "flat", and "sprat" because they share "at" with the rime of the "cat." The rime cognate also contains multisyllabic words such as "matins" having "at" in the first syllable, and "cravat" having "at" in the second syllable.

In the present study, the three sets of word candidates, the neighborhood, the cohort, and the rime cognate, are compared for their reliability as a lexical competitor set in spoken word recognition. Three sets of analyses were carried out. The first analysis showed that the rime cognate is better than the neighborhood and the cohort in terms of correlations with word recognition rates. The second and third analyses showed that the rime cognate includes the reliable part of the neighborhood and the cohort, which indicate that the rime cognate integrates the useful portions of the neighborhood and the cohort.

Analysis 1

This analysis addresses the question of which set of word candidates is the best predictor for lexical competition in spoken word recognition among the neighborhood, cohort, and rime cognate. The best candidate set should have the greatest negative correlations to word recognition rate in terms of the density, the mean frequency, and/or the maximum frequency, as a reflections of lexical competition.

Materials

Analysis 1 was conducted on the recognition rates of English spoken words which were extracted from the "Easy-Hard" word multi-talker speech database (Torretta, 1995). The database contains recognition rates of 150 monosyllabic words pronounced by 10 talkers at three speaking rates of slow, medium, and fast. The 150 words pronounced by each talker at each speaking rate were presented to 10 subjects to obtain a percent correct identification score. A total of 300 subjects (10 subjects X 10 talkers X 3 rates) participated in supplying responses for the database.

Three items out of the 150 words were not used for the analysis. They were "white" and two "wrong"s. It is because the "white" might be a CCVC word rather than a CVC word, and because there are two entries for the "wrong" in the database. After excluding these three items, the recognition rates of the 147 CVC words (See Appendix) in three speaking rates were used for Analysis 1. There were 441 items in total. Word recognition rates for these items were obtained by averaging the identification score over ten talkers.

Definition of Candidate Set

Definitions of the candidate sets used for Analysis 1 are shown in Table 1. "Neighborhood" is defined as a union of three set of words of the first, the second, and the third phoneme substitution with a target word (Pisoni, Nusbaum, Luce, & Slowiaczek, 1985; Frauenfelder, 1990; Frauenfelder, Baayen, Hellwig, & Schreuder, 1993).

"Cohort" is defined as a set of words which share the first two phonemes (i.e., CV) with a target CVC word. This definition is consistent with the word-initial cohort which was used in the studies of Bard and Shillcock (1993) and Marslen-Wilson, Moss, and van Halen (1996).

"Rime Cognate" is defined as a set of words which have the same phoneme sequence with the rime (i.e., VC) of a target CVC word. The rime cognate used in Analysis 1 does not include the words which have the same phoneme sequence with the rime of the target CVC word at the second syllable or at the later syllables. For example, the rime cognate of the target word of "cat" does not include "cravat" which has the same phoneme sequence with the rime at the second syllable. This was done because it is hard to imagine that such word candidates are activated for a target word presented in isolation. Because there is enough silence before the target word, it clearly indicates the beginning of the target word and it removes the possibilities for the activation of such word candidates. However, if the target word is embedded in continuous speech, it would be reasonable to include such word candidates, because the beginning of the word is sometimes ambiguous in the continuous speech.

Table 1.

**Phoneme sequence pattern of word candidates
in neighborhood, cohort, and rime cognate for CVC target word.**

Phoneme Sequence Pattern

Neighborhood	?VC C?C CV?
Cohort	CV(*)
Rime Cognate	(c+)VC(*)

Note.

C = Common consonant to target word,
 V = Common vowel to target word,
 ? = Any one phoneme,
 * = Any length of phoneme sequence,
 c+ = Any length of consonant sequence.
 Items in parenthesis can be null.

Method

Density, mean log frequency, and the maximum log frequency were obtained for the neighborhood, the cohort, and the rime cognate of the 147 target words using a computerized dictionary (Nusbaum, Pisoni, & Davis, 1984) which contains 19,295 words with Kucera and Francis (1967) word count. Log frequency of the 147 target words was also obtained from the dictionary. Summary statistics for these variables are shown in Table 2.

By excluding the factor of log frequency of the target word, Spearman's rank order partial correlation coefficient was calculated between the word recognition rate and the three variables of density, mean log frequency, and maximum log frequency. The Spearman's rank order partial correlation coefficient is suitable to the analysis because there is a weak tendency that the density and the frequency of the neighborhood are higher for high-frequency target words than for low-frequency target words (Frauenfelder, Baayen, Hellwig, & Schreuder, 1993; Pisoni, Nusbaum, Luce, & Slowiaczek, 1985), and because the distributions are skewed for the density, mean log frequency and maximum log frequency of the neighborhood (Frauenfelder, Baayen, Hellwig, & Schreuder, 1993) and cohort (Bard & Shillcock, 1993) as well as for the log frequency of a target word.

Table 2.

Statistics of the CVC words for Analysis 1 (N=147).

	Mean	SD	Minimum	Maximum
Target Word				
Log Frequency	1.39	0.83	0.30	3.99
Neighborhood				
Density	17.56	6.77	1	31
Mean Log Frequency	1.84	0.48	0.56	2.89
Maximum Log Frequency	2.69	0.70	0.85	4.03
Cohort				
Density	56.27	61.80	1	325
Mean Log Frequency	1.34	0.48	0.30	2.89
Maximum Log Frequency	2.43	0.79	0.30	4.42
Rime Cognate				
Density	109.00	143.60	1	973
Mean Log Frequency	1.34	0.37	0.30	2.27
Maximum Log Frequency	2.64	0.77	0.30	4.46

Results

The Spearman's rank order partial correlation coefficients are shown in Figure 1. The results are shown for the data of all speaking rates in Figure 1, because no substantial differences were observed among the results of the slow, medium, and fast speaking rate.

The partial correlation coefficient significantly differed from zero for all variables of the rime cognate; density ($r = -.293, p < .01$), mean log frequency ($r = -.145, p < .01$), and maximum frequency ($r = -.215, p < .01$). The partial correlation coefficient also differed from zero for the density of the neighborhood ($r = -.192, p < .01$), and the density of the cohort ($r = -.104, p < .05$). However, it was not significantly different from zero for the mean log frequency and the maximum frequency of the neighborhood and cohort.

For the density, the partial correlation coefficient was different between the rime cognate and the neighborhood ($p < .05$), and between the rime cognate and the cohort ($p < .01$). For the mean log frequency, the partial correlation coefficient was different between the rime cognate and the cohort ($p < .01$), and between the cohort and the neighborhood ($p < .05$). For the maximum log frequency, the partial correlation coefficient was different between the rime cognate and the neighborhood ($p < .05$), and between the rime cognate and the cohort ($p < .01$).

For the rime cognate, the partial correlation coefficient was different between the density and the mean log frequency ($p < .05$), and between the mean log frequency and the maximum log frequency ($p < .01$). However, it was not significantly different between the density and the maximum log frequency. For the neighborhood, the partial correlation coefficient was different between the density and the mean log frequency ($p < .05$), and between the density and the maximum log frequency ($p < .05$). For the cohort, the

partial correlation coefficient was different between the density and the mean log frequency ($p < .05$), and between the density and the maximum log frequency ($p < .05$).

Insert Figure 1 about here.

Discussion

Results of this first analysis clearly show that the rime cognate is a better predictor of lexical competition in spoken word recognition than the neighborhood and the cohort. The partial correlation coefficient has a larger negative value for the rime cognate than the neighborhood and the cohort in terms of the density, the mean log frequency, and the maximum log frequency.

All partial correlation coefficients were negative for the rime cognate. This pattern indicates that the set of words which have the same phoneme sequence with the rime (i.e., VC) of a CVC target word are activated and they compete against the target word. Notice that the negative effects are not an artifact of the target frequency, because the contribution of the target frequency is excluded by calculating partial correlations. Hence, the effects are purely based on the structural characteristics of the rime cognate.

For the rime cognate, the density had the greatest contribution to word recognition, the maximum log frequency was the second, and the mean log frequency was the least. Although there were significant differences between the density and the mean log frequency, there were no significant differences between the density and the maximum log frequency. This fact is ambiguous in terms of the number of activated competitors. Significant partial correlation coefficient for the density indicates that all of the competitors in the rime cognate affect the word recognition. On the other hand, significant partial correlation coefficient for the maximum log frequency indicates that only one competitor with the highest frequency in the rime cognate affects the word recognition.

The significant partial correlation coefficient for the maximum log frequency might be an artifact of the effect of mean log frequency. If the maximum log frequency has a high positive correlation with the mean log frequency, it automatically produces significant relationship to word recognition when the mean log frequency is significantly related to word recognition. And, this appears to be the case for the current data set. Spearman's rank order partial correlation coefficient was very high between the mean log frequency and the maximum log frequency ($r = .844$, $p < .001$). Therefore, the pattern is consistent. All the competitors in the rime cognate affect the word recognition, but one competitor with the highest frequency does not.

How does both the density and the frequency of the rime cognate affect word recognition? One possibility is that the amount of effectiveness of each competitor is modulated by its frequency. That is, a high frequency competitor is more effective than low frequency competitor in the rime cognate.

Bard (1990) and Bard and Shillcock (1993) have already mentioned this possibility of frequency modulation of competitors. And, it has been incorporated in some models of spoken word recognition. For example, TRACE model (McClelland & Elman, 1986) achieves such modulation by changing the resting activation level of a word unit.

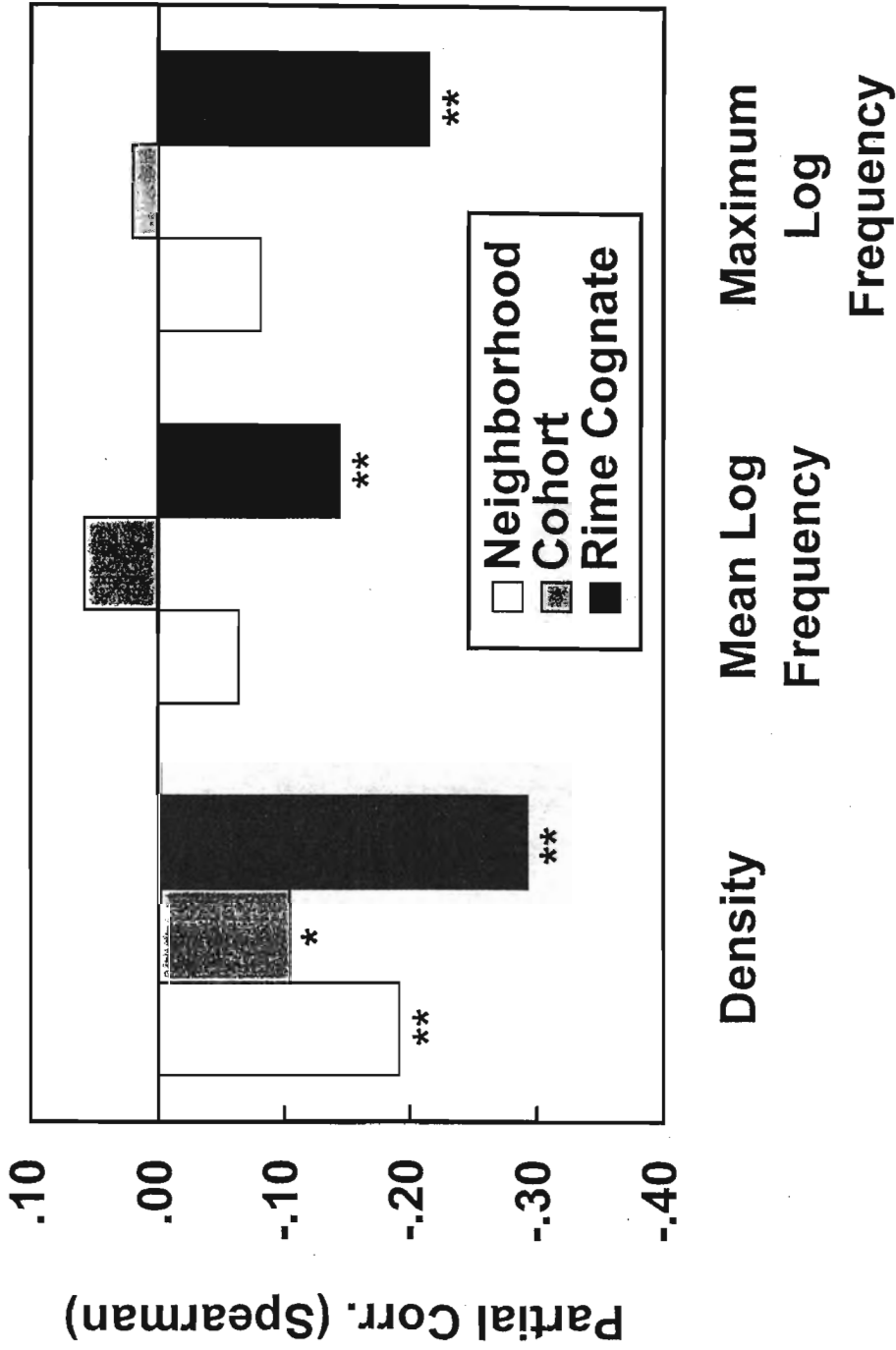


Figure 1. Spearman's rank order partial correlation coefficient between word recognition rate and density, mean log frequency, and maximum log frequency of neighborhood, cohort, and rime cognate. Excluding factor for the partial correlation is log frequency of target word. Single and double asterisks respectively represent difference from zero at 5% and 1% significance level.

However, Bard (1990) and Bard and Shillcock (1993) analyzed the cohort only. And, the TRACE model basically uses the cohort-type candidate set. Further research, therefore, is necessary to reveal the characteristics of the frequency modulation in the rime cognate.

Analysis 2

Although Analysis 1 showed that the rime cognate is reliable on spoken word recognition in terms of partial correlation with identification performance, many studies have shown that the neighborhood is also reliable. Analysis 2 addresses the question of why the neighborhood is reliable in some degree.

The conventional view of a lexical neighborhood is a union of positional neighborhoods in which a phoneme is substituted to a target word at one of phoneme positions (e.g., Frauenfelder, 1990; Frauenfelder, Baayen, Hellwig, & Schreuder, 1993; Pisoni, Nusbaum, Luce & Slowiaczek, 1985). In case of the CVC words, the conventional neighborhood consists of the positional neighborhoods at the first (?VC), the second (C?C), and the third phoneme position (?VC) as shown in Table 1. Among these positional neighborhoods, the positional neighborhood with the first phoneme substitution (?VC) is included as a part of the rime cognate ((c+)VC(*)), because it shares the second and the third phoneme (i.e., the rime) with the target word as the rime cognate does.

It is possible that the conventional neighborhood is reliable because this positional neighborhood contributes much to the word recognition. More specifically, the hypothesis is that a positional neighborhood with the first phoneme substitution would be more reliable than other positional neighborhoods with the second or the third phoneme substitution. Analysis 2 assesses this hypothesis.

Materials

Word recognition rates used in Analysis 1 were also used in Analysis 2.

Definition of Candidate Set

Analysis 2 was conducted on the positional neighborhoods of which a phoneme was substituted to the target word at one of phoneme positions. The positional neighborhood of the first phoneme position is the set of words which share the second phoneme and the third phoneme with the target word but not the first phoneme (?VC). The positional neighborhood of the second phoneme position is the set of words which share the first phoneme and the third phoneme with the target word but not the second phoneme (C?C). The positional neighborhood of the third phoneme position is the set of words which share the first phoneme and the second phoneme with the target word but not the third phoneme (CV?).

Method

For each positional neighborhood of the 147 target words, density, mean log frequency, and the maximum log frequency were obtained from the same computerized dictionary as Analysis 1. Summary statistics for these variables are shown in Table 3.

By excluding the factor of log frequency of a target word, Spearman's rank order partial correlation coefficient was calculated between word recognition rates and the three variables of density, mean log frequency, and maximum log frequency for each neighborhood.

Table 3.

Statistics of the positional neighborhood of the CVC words for Analysis 2 (N=147).

Positional Neighborhood	Mean	SD	Minimum	Maximum
1st Phoneme Position				
Density	7.03	3.86	0	16
Mean Log Frequency	1.59	0.73	2.97	0.39
Maximum Log Frequency	2.13	0.96	4.03	0.63
2nd Phoneme Position				
Density	4.76	2.82	0	10
Mean Log Frequency	1.49	0.74	2.92	0.46
Maximum Log Frequency	1.89	0.95	3.71	0.72
3rd Phoneme Position				
Density	5.78	2.94	0	13
Mean Log Frequency	1.52	0.71	3.09	0.44
Maximum Log Frequency	1.99	0.89	3.86	0.67

Results

Figure 2 shows Spearman’s rank order partial correlation coefficients between word recognition rate and the variables of the positional neighborhoods. Figure 2 also shows the partial correlation obtained in Analysis 1 for the conventional neighborhood which is represented by “U” (for “union”).

For the positional neighborhood of the first phoneme position (?VC), the partial correlation coefficient significantly differed from zero for the density ($r = -.158, p < .01$), the mean log frequency ($r = -.213, p < .01$), and the maximum log frequency ($r = -.216, p < .01$). For the positional neighborhood of the second phoneme position (C?C), the partial correlation coefficient significantly differed from zero only for the density ($r = -.155, p < .01$). For the conventional neighborhood (U), the partial correlation coefficient significantly differed from zero for the density ($r = -.192, p < .01$) as shown in Analysis 1. No significant differences were observed between any pair of these significant partial correlation coefficients.

For the density, the partial correlation coefficient was significantly different between the conventional neighborhood (U) and the positional neighborhood of the third phoneme position (CV?) ($p < .01$).

For the mean log frequency, the partial correlation coefficient for the positional neighborhood of the first phoneme position (?VC) was significantly different from that for the conventional neighborhood (U) ($p < .01$), for the positional neighborhood of the second phoneme position (C?C) ($p < .01$), and for the positional neighborhood of the third phoneme position (CV?) ($p < .05$).

For the maximum log frequency, the partial correlation coefficient for the positional neighborhood of the first phoneme position (?VC) was significantly different from that for the conventional neighborhood (U) ($p < .01$), for the positional neighborhood of the second phoneme position (C?C) ($p < .01$), and for the positional neighborhood of the third phoneme position (CV?) ($p < .05$).

 Insert Figure 2 about here.

Discussion

Although the positional neighborhood of the first phoneme substitution (?VC) was not different for the positional neighborhood of the second phoneme substitution (C?C) in terms of the density, it showed a much larger negative partial correlation coefficient than other positional neighborhoods in terms of the density, the mean log frequency, and the maximum log frequency. These results indicate that the positional neighborhood with the first phoneme substitution (?VC) is more reliable than other positional neighborhoods with the second (C?C) or the third phoneme substitution (CV?).

The positional neighborhood of the first phoneme substitution (?VC) is a part of the conventional neighborhood (U) which has reliable effects of density as shown in Analysis 1. Spearman's rank order partial correlation coefficients between these two neighborhoods was positive and high ($r = .691$) in terms of the density. This pattern means that the conventional neighborhood is reliable because it contains the positional neighborhood of the first phoneme substitution which is reliable.

The positional neighborhood of the first phoneme substitution is included in the rime cognate which is the best candidate set in Analysis 1. Therefore, the results of Analysis 2 suggest that the conventional neighborhood is reliable in some degree because it shares a part with the rime cognate which is more appropriate as a word candidate set.

Analysis 3

In Analysis 1, the cohort was defined as a set of words which share the first two phonemes (i.e., CV part) with a target CVC word. However, cohort can be defined differently because the number of shared phonemes is not explicitly specified for the word-initial cohort. The word-initial cohort can share one or two "segments" at word onset (Marslen-Wilson, 1989) or initial "sequence" (Marslen-Wilson, 1984) which correspond to 100 to 150 ms of speech waveform (Marslen-Wilson, 1987). Therefore, the cohort can be defined as a set of words sharing only an initial phoneme, first two phonemes, or first three phonemes (Marslen-Wilson & Welsh, 1978). These three kind of sets were used in Analysis 3.

In case of the CVC words, the cohort sharing the first three phonemes (i.e., all phonemes) with the target word is completely included in the rime cognate. On the other hand, the cohorts sharing the first phoneme or the first two phonemes are only partially included. Notice that the candidates which are common with the rime cognate are the same across the these cohorts, because the cohorts sharing the first phoneme and the first two phonemes include the cohort sharing the first three phonemes.

However, the proportion of the common candidates with the rime cognate is larger in the cohort sharing the first three phonemes than in other cohorts. In other words, the cohorts sharing the first phoneme or the first two phonemes have larger numbers of uncommon candidates with the rime cognate than the cohort sharing the first three phonemes.

Therefore, it is likely that the cohort sharing the first three phonemes might be more reliable in word recognition than the other cohorts, because the former cohort is a part of the rime cognate which is reliable in the word recognition. This possibility was examined in Analysis 3.

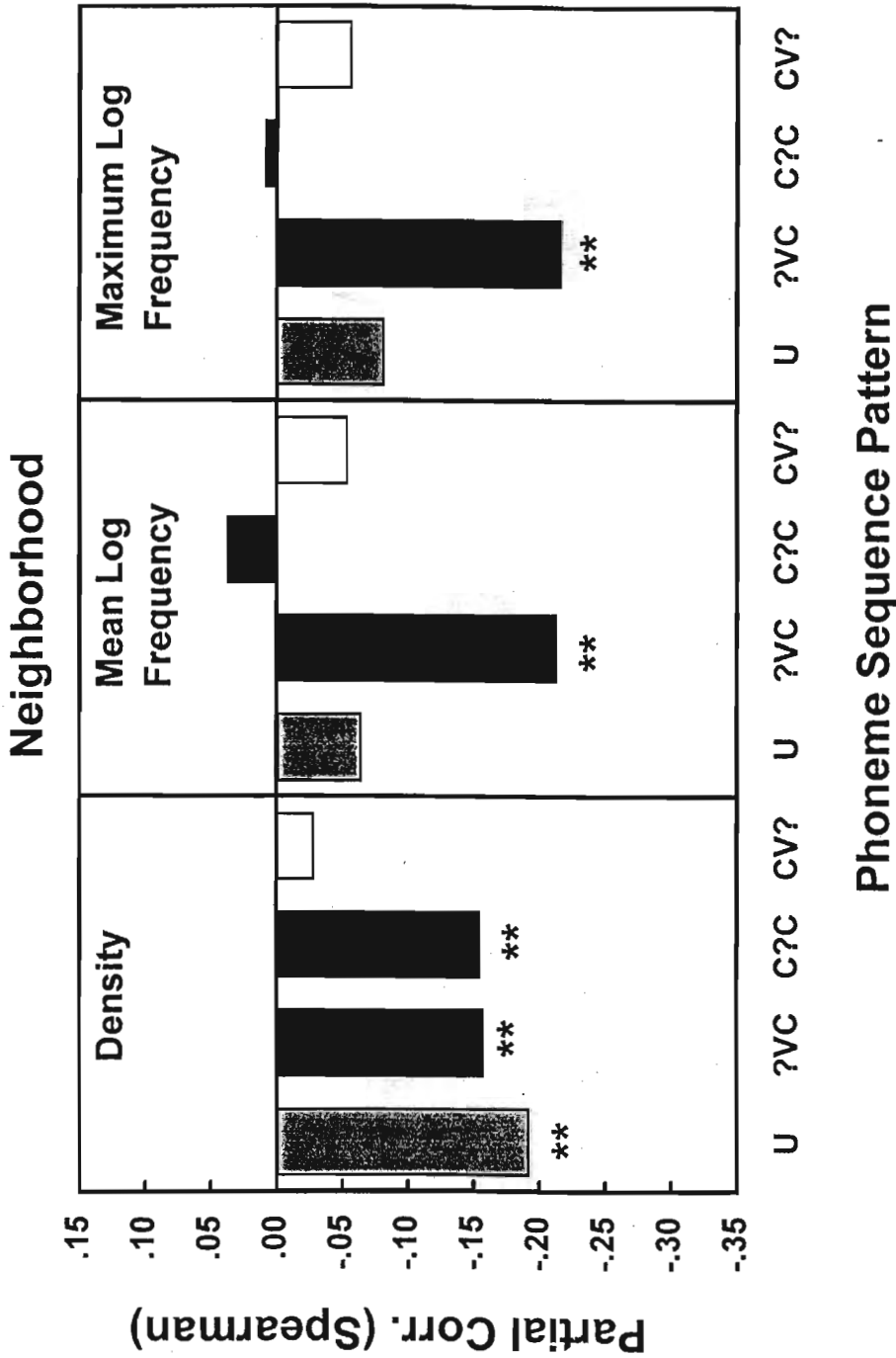


Figure 2. Spearman's rank order partial correlation coefficient between word recognition rate and density, mean log frequency, and maximum log frequency of neighborhood. Excluding factor for the partial correlation is log frequency of target word. "?" represents a substitutable single phoneme. "C" and "V" respectively represent a common consonant and vowel with the CVC target word. "U" represents the conventional neighborhood. Double asterisks show the significant difference from zero ($p < .01$).

Materials

Word recognition rates used in Analysis 1 were also used in Analysis 3.

Definition of Candidate Set

Three cohorts were used in Analysis 3. They were the cohorts sharing the first phoneme, the first two phonemes, and the first three phonemes. The cohort sharing the first phoneme is represented as C(*) in which the common first phoneme with a target word is followed by any sequence of phonemes. The cohort sharing the first two phonemes is represented as CV(*), which is the same one as in Analysis 1. The cohort sharing the first three phonemes is represented as CVC(*) in which an entire phoneme sequence of the target word is followed by any sequence of phonemes.

Method

For each cohort of the 147 target words, density, mean log frequency, and the maximum log frequency were obtained from the same computerized dictionary as Analysis 1. Statistics of these variables are shown in Table 4.

By excluding the factor of log frequency of a target word, Spearman's rank order partial correlation coefficient was calculated between the word recognition rate and the three variables of density, mean log frequency, and maximum log frequency.

Table 4.

Statistics of the cohort of the CVC words for Analysis 3 (N=147).

Cohort	Mean	SD	Minimum	Maximum
One Phoneme Sharing (C(*))				
Density	902.50	520.07	116	1895
Mean Log Frequency	1.46	0.27	1.00	2.09
Maximum Log Frequency	3.36	0.46	2.55	4.42
Two Phoneme Sharing (CV(*))				
Density	56.27	61.80	1	325
Mean Log Frequency	1.34	0.48	0.30	2.89
Maximum Log Frequency	2.43	0.79	0.30	4.42
Three Phoneme Sharing (CVC(*))				
Density	4.65	10.92	0	125
Mean Log Frequency	0.61	0.53	0	2.64
Maximum Log Frequency	0.83	0.73	0	2.64

Results

Figure 3 shows Spearman's rank order partial correlation coefficient. The results of Analysis 1 for the cohort sharing the first two phonemes were replotted as the notation with CV(*) in Figure 3.

For the cohort sharing the first three phonemes (CVC(*)), the partial correlation coefficient significantly differed from zero for the density ($r = -.232$, $p < .01$), the mean log frequency ($r = -.147$, $p < .01$), and the maximum log frequency ($r = -.172$, $p < .01$). For the cohort sharing the first two phonemes (CV(*)), the partial correlation coefficient significantly differed from zero only for the density ($r = -.104$, $p < .05$), as already shown in Analysis 1.

For the cohort sharing the first three phonemes (CVC(*)), the partial correlation coefficient was significantly different between the density and the mean log frequency ($p < .05$), and between the mean log frequency and the maximum log frequency ($p < .01$). But there was no difference of the partial correlation coefficient between the density and the maximum log frequency.

For the density, the partial correlation coefficient for the cohort sharing the first three phonemes (CVC(*)) was significantly different from that for the cohort sharing the first two phonemes (CV(*)) ($p < .01$), and for the cohort sharing the first phoneme (C(*)) ($p < .01$).

For the mean log frequency, the partial correlation coefficient for the cohort sharing the first three phonemes (CVC(*)) was significantly different from that for the cohort sharing the first two phonemes (CV(*)) ($p < .01$), and for the cohort sharing the first phoneme (C(*)) ($p < .05$).

For the maximum log frequency, the partial correlation coefficient for the cohort sharing the first three phonemes (CVC(*)) was significantly different from that for the cohort sharing the first two phonemes (CV(*)) ($p < .01$), and for the cohort sharing the first phoneme (C(*)) ($p < .01$).

 Insert Figure 3 about here.

Discussion

Analysis 3 shows that the cohort sharing the first three phonemes (CVC(*)) is reliable in predicting word recognition performance in terms of the density, the mean log frequency, and the maximum log frequency. On the other hand, the cohorts sharing the first phoneme (C(*)) was not reliable in all cases. And the cohort sharing the first two phonemes (CV(*)) is reliable in terms of the density, but the partial correlation coefficient is much smaller than that found for the cohort sharing the first three phonemes.

These findings indicate that the cohort is reliable to the word recognition when all or a large part of it is included in the rime cognate. The present findings suggest that the cohort sharing the first two phonemes (CV*), which is usually used as the definition of the cohort, is in some degree reliable for the word recognition, because it contains the most distinctive part cohort (CVC*) which is included in the rime cognate. In other words, the reliability of the cohort in spoken word recognition derives from that of the rime cognate.

General Discussion

Analysis 1 showed that the rime cognate is a good predictor for lexical competition in spoken word recognition. It has much larger negative correlation than the neighborhood and the cohort. Analysis 2 and 3 showed that the rime cognate includes the parts of the neighborhood and the cohort which have greater negative correlation with recognition performance than other parts of them. In other words, the rime

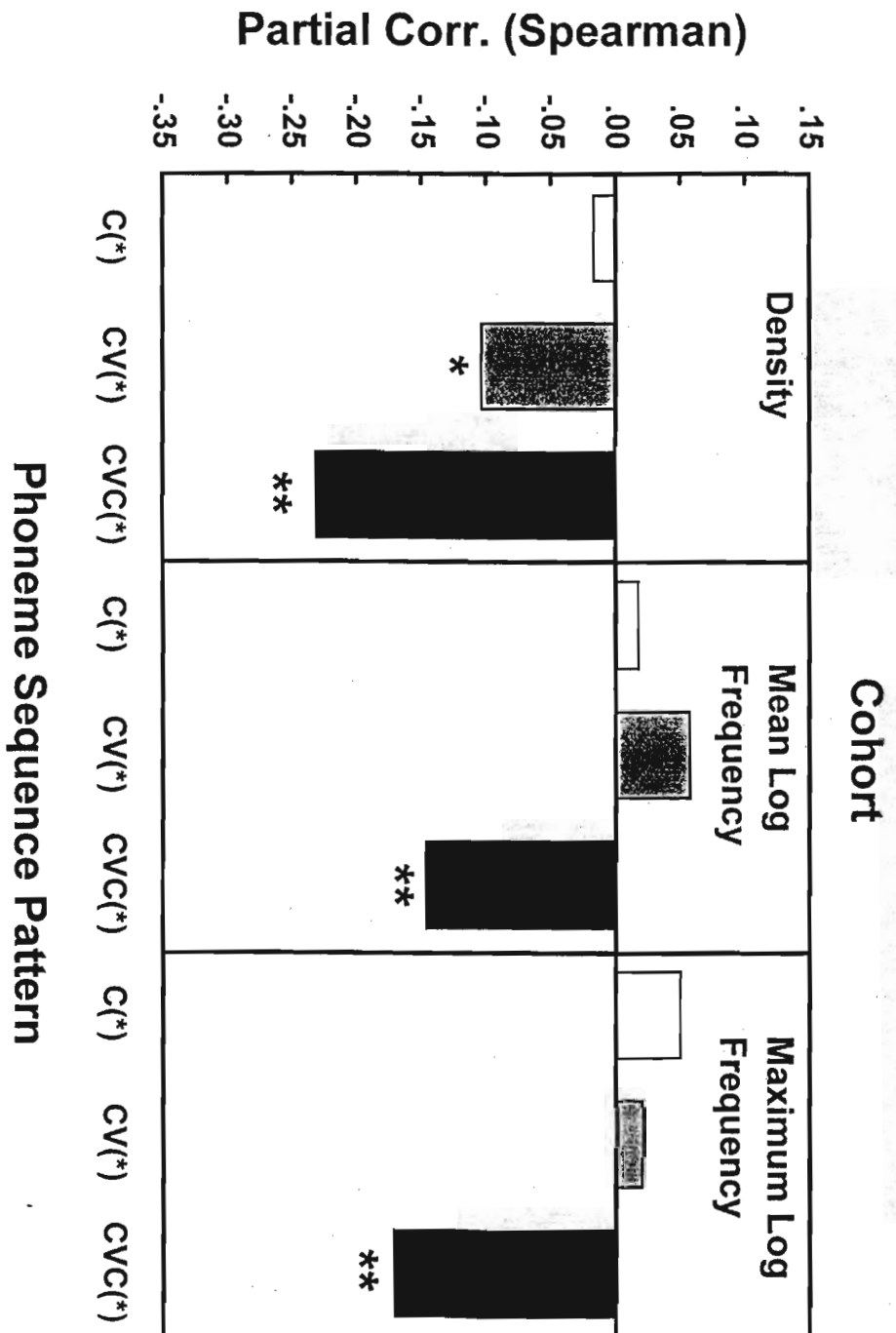


Figure 3. Spearman's rank order partial correlation coefficient between word recognition rate and density, mean log frequency, and maximum log frequency of cohort. Excluding factor for the partial correlation is log frequency of target word. "*" represents any length of phoneme sequence including null, "C" and "V" respectively represent a common consonant and vowel with the CVC target word. Single and double asterisks respectively represent difference from zero at 5% and 1% significance level.

cognate integrates the neighborhood and the cohort by discarding their unimportant parts. The integration provided by the rime cognate dissolves the discrepancy between the neighborhood and the cohort in the previous studies.

Strictly speaking, the reliability of the rime cognate is only shown for the CVC English words in this study. Thus, there are several unsolved questions. For example, is the rime cognate reliable for the words which do not have the CVC phoneme sequence pattern? In other words, monosyllabic words with more than two phonemes in the onset and/or the coda have not been investigated from the view point of the rime cognate. Multisyllabic words have not been investigated either.

However, there is some indirect support for that the rime cognate is a reliable word candidate set for non-CVC monosyllabic words. This comes from studies of phonological priming. Slowiaczek and Hamburger (1992) have shown that the prime word produces interference and increases reaction times of shadowing of the target word when there is an overlap of the initial three phonemes between the prime and the target word. In contrast, the prime word has no effects when the overlap is the first two phonemes, and it produces facilitative effects when the overlap is the initial phoneme.

Slowiaczek and Hamburger (1992) used monosyllables with four or five phonemes in length in their experiments. Phoneme sequence pattern of their stimuli and their proportion were CVCC (30%), CCVC (58%), CCVCC (11%), and CCCVC (1%) (L.M. Slowiaczek, personal communication, July 1, 1997).

Three phoneme overlap includes a part of the rime. That is, the vowel nucleus and the first phoneme of the coda for in CVCC pattern, and the vowel nucleus for CCVC and CCVCC pattern. The proportion of such stimuli was 99% in total. On other hand, two phoneme overlap only includes the vowel nucleus in CVCC pattern which is only 30% of the stimuli. And, one phoneme overlap did not include any part of the rime. Therefore, the possibility is larger for three phoneme overlap to include a part of the rime than one or two phoneme overlap.

Slowiaczek and Hamburger (Hamburger & Slowiaczek, 1996, Slowiaczek & Hamburger, 1992) argued that the interference in the three phoneme overlap reflects lexical competition, but the facilitative effect in the one phoneme overlap reflects prelexical processing. The interference in the three phoneme overlap can be interpreted as support for the proposal that the rime cognate is activated by a part of the rime and lexical competition among the rime cognate causes negative effects on phonological priming. However, because the rime cognate is activated by a part of the rime, not the whole rime, by this interpretation, it is a little bit different from the original definition of the rime cognate. Therefore, their results only indirectly suggest that the rime cognate is a reliable competitor set for the non-CVC monosyllabic words.

Some indirect support for reliability of the rime cognate on multisyllabic words comes from studies of word spotting. McQueen, Norris, and Cutler (1994) have shown that the multisyllabic word (domestic) which shares its beginning part with the carrier nonword (domes) competes with the monosyllabic target word (mess). Their results suggest that the monosyllabic target word and the multisyllabic word compete with each other because they are related with the rime. For example, "mess" and the second syllable of "domestic" share the rime of /es/. Hence, their results suggest that the rime cognate is activated at the second syllable in a multisyllabic word.

A study by Treiman, Fowler, Gross, Berch, and Weatherston (1995) provides additional support for the rime cognate in recognition of multisyllabic words. They conducted a series of experiments using word games in which a subject produced a new nonword by changing one or two phonemes in the noninitial syllable in a disyllable and trisyllable multisyllabic nonword. They found that subjects performed better in the games in which a changing unit corresponded to the onset and the rime than the games in which the changing unit does not correspond to the onset and the rime. Although their stimuli are nonwords, it is highly probable that the units of the rime and the onset also affect the recognition of a word with multisyllable structure. These findings suggest that the rime cognate may also be a reliable candidate set for multisyllabic words as well.

Further support for this proposal comes from the studies of the neighborhood of multisyllabic words. Cluff and Luce (1990) obtained recognition scores of two-syllable spondee words (e.g., "icecream", horseshoe). They manipulated the frequency, the neighborhood density, and the neighborhood frequency for each syllable in the word and found that the recognition scores of the two-syllable words are affected by these variables for each syllable. As shown in Analysis 2, a reliable part of neighborhood is included in the rime cognate. Therefore, if the neighborhood of each syllable affects the recognition of multisyllabic words, it is probable that the rime cognate of each syllable also affects it, too.

Taken together, the overall pattern suggests that spoken word recognition is affected by the rime cognate of non-CVC mono-/multi-syllabic words as well as the CVC monosyllabic words.

Another question that can be addressed here concerns the metrical stress and the rime cognate in multisyllabic words. Is the rime cognate activated by the rime of strong syllable, weak syllable, or both? The answer is probably that it is only activated by the rime of the strong syllable, because indirect evidence for the rime cognate has been obtained only in the strong syllable so far. For example, in McQueen, Norris, and Cutler (1994)'s study, competition was observed between "domestic" and "mess" which corresponds to the strong syllable of the "domestic." In Cluff and Luce's (1990) study, the spondees by definition consist of two strong syllables.

In addition, the strong syllable is more informative than the weak syllable because the number of vowel nuclei is larger in the strong syllable than in the weak syllable. These characteristics suggest that the weak syllable produces a larger rime cognate than the strong syllable. In other words, the weak syllable is not useful to reduce the number of word candidates in the rime cognate but the strong syllable does. Therefore, it is reasonable that the rime cognate is only activated by the rime of the strong syllable.

If the rime cognate is activated by the rime of the strong syllable with the rime matching strategy (RMS), it is important to consider the relationship to the metrical segmentation strategy (MSS) proposed by Cutler and Norris (1988). Although the MSS is used for prelexical processing, the MSS interacts with lexical processing (McQueen, Norris, & Cutler, 1994). The MSS is triggered by the beginning of a strong syllable. This is not consistent with the RMS which is triggered by the rime of the strong syllable. However, the MSS can be consistent to the RMS with a small modification. That is, the MSS can be modified to be triggered by the rime of the strong syllable. The justification for this modification is that there is not enough information in the consonant cluster of the onset to determine whether the syllable is strong or weak, and that the strong syllable is detected only after the vowel nucleus is processed. This modification permits the MSS be more realistic in the left-to-right processing and also be consistent to the RMS.

An example of this modification is shown as following. The recent version of SHORTLIST model (Norris, McQueen, & Culter, 1995) simulates the spoken word recognition incorporating the MSS by boosting the activation level of word units which have the strong syllable. In their simulation, the word candidates are boosted at the beginning of the strong syllable (i.e., at the onset of the strong syllable). However, if the modified MSS is applied, the word candidates should be boosted at the beginning of the rime of the strong syllable. In this case, boosted word candidates by the MSS will coincide to the rime cognate by the RMS.

Another question that needs to be addressed is the language dependency of the rime cognate. Is the rime cognate reliable in other languages? The answer might be negative. The rime cognate is probably reliable only in English (and some other language which use stress), because it may be activated by the rime of strong syllable. The strong syllable is a processing unit for segmentation in English (Cutler & Norris, 1988), but it is not for other languages. Different processing units may be in other languages. For example, it is a mora for Japanese (Otake, Hatano, Cutler, & Mehler, 1993; Cutler & Otake, 1994), and a syllable for French (Mehler, Dommergues, Frauenfelder, & Segui, 1981) and also for Spanish and Catalan (Sebastian-Galles, Dupoux, Segui, & Mehler, 1992). Although a specific processing unit for segmentation does not necessarily coincide to a lexical access unit, they are probably identical because of parsimony of processing. Therefore, the rime cognate which may depend on the strong syllable might not be reliable in other languages which do not use the strong syllable as a processing unit.

More direct negative evidence for the rime cognate has been obtained for Japanese. Watanabe (1996) measured the reaction times for discrimination between a target word and nontarget words which diverge from the target word at a certain phoneme in sequential matching. Using a regression equation for phoneme detection time (Amano, 1995), he found that the reaction times for the nontarget words have almost constant delay from detection time of the divergent phoneme at which the nontarget word can be discriminated from the target word in left-to-right matching procedure. His findings coincide with the original cohort model (Marslen-Wilson & Welsh, 1978) which claims that a word is processed in sequential matching and is perceived at the uniqueness point where the word diverges from all other word candidates. Although the divergent point in Watanabe's study does not correspond to the actual uniqueness point in a mental lexicon, his findings strongly suggest that the cohort, but not the rime cognate, is a reliable candidate set for recognition of Japanese spoken words.

Other studies implicitly assume that lexical access is universal among languages. However, it is highly probable that the lexical access is language dependent like segmentation of speech depends on language.

Based on the results and some considerations in this study, rime cognate model (RCM) is proposed for spoken word recognition as a set of working hypotheses. The RCM has following characteristics.

1. The rime of a given word is used as a cue for lexical access. Word candidates sharing the same phoneme sequence with the rime of a given word are activated when the rime is processed. This lexical matching procedure is called rime matching strategy (RMS). The set of word candidates which is activated by the RMS is called the rime cognate. The rime cognate is activated only by the rime in a strong syllable but not by the rime in a weak syllable. The rime of given input word is matched to the phoneme sequence in every possible positions within a word in a mental lexicon.

2. The word candidates in the rime cognate compete against each other. Competition is realized by inhibitory information flow among word candidates. Competition continues until one candidate is highly activated and it exceeds a threshold for recognition.

3. Phonemes which are not in the rime of a strong syllable are used for increasing or decreasing the activation of word candidates. That is, facilitatory and inhibitory information are sent to word candidates from phonemes according to their consistency to words. This bottom-up information flow begins only after the rime cognate is developed by the RMS. Before that, no information flow exists from the phonemes to word candidates. The inhibitory information flow from phonemes to word candidates might be redundant if word candidates have inhibitory connections with each other. However, it may facilitate word recognition because irrelevant word candidates are deactivated sooner, which allows a relevant word candidate to be a winner more quickly.

4. A word candidate facilitates phoneme processing by changing phoneme's gain of activation. The word candidate makes the gain of the phoneme be greater in proportion to activation level of the word candidate when the phoneme is included in the word. These characteristics suggest that the facilitation on phoneme processing should be effective only after the rime cognate is developed, because the word candidate is not activated before it. And, the phoneme must have received some bottom-up information for being facilitated by the word, because the word only changed the gain of the phoneme but not the activation level itself.

5. Gain of activation of a word candidate is modulated by its word frequency and familiarity. A high-frequency/high-familiarity word has greater gain than a low-frequency/low-familiarity word, so that the former is more easily activated than the latter.

6. Syntactic and semantic information create a bias on the rime cognate by modifying the gain of word candidates. That is, syntactically and semantically correct word candidates are more easily activated than incorrect word candidates. This means that syntactic and semantic information indirectly reduces the number of word candidates after the lexical access is conducted by the RMS.

Some of the characteristics of the RCM may be easily implemented in the TRACE model (McClelland & Elman, 1986) and the SHORTLIST model (Norris, 1994; Norris, McQueen, & Cutler, 1995). For example, the RMS can be achieved by putting larger weights on the connections between the word units and the phoneme units which correspond to the rime. However, to precisely simulate the activation of a set for the rime cognate, they should have a gating mechanism which blocks information flow from phoneme units in the onset to word units until the rime is processed. This gating mechanism might be achieved by putting extremely large weights on the rime part and very small weights on the onset part. Then, word units which are slightly activated by the onset virtually would have null activation when word units are activated by the rime.

Other characteristics of the RCM might require a fundamental changes in the TRACE model and the SHORTLIST model. For example, the inhibitory connections between a phoneme unit and a word unit have not been implemented in either model. Effects of introducing inhibitory connection are unknown for the model's performance. However, the SHORTLIST model implicitly has such inhibitory connections by limiting the number of word candidates in word level.

The gain control might cause a fundamental change for these models, because they do not have such function. However, the gain control by top-down information has some merits which the activation

control does not have in these models. That is, the top-down information affects a unit only if the unit receives some bottom-up information. This means bottom-up information dominates speech recognition. If top-down information affects a unit via adjustments in activation levels, the unit (i.e., a phoneme or a word) can be recognized without any bottom-up information when the top-down information is strong enough to fully activate the unit. This means that words can be recognized without speech input. The gain control can avoid such unrealistic situation.

The proposed characteristics of the RCM may have some superiority over the TRACE model and the SHORTLIST model. However, the set of hypotheses have not been confirmed. Further research is necessary to examine the predictions of the RCM.

In summary, a set of analyses were carried out using rime cognate as a candidate set for lexical processing in spoken word recognition. It is revealed that, at least for CVC English words, the rime cognate is more reliable lexical competitor set than the lexical neighborhood and the word-initial cohort, and that the rime cognate integrates the neighborhood and cohort by including their reliable part. It is suggested that the rime cognate is also a reliable candidate set for non-CVC words and multisyllabic words. But the rime cognate may be reliable to only spoken word recognition in English.

References

- Amano, S. (1995). Time-course of phoneme/word perception in Japanese (*Technical Report of Hearing, H-95-50*). Tokyo: Acoustical Society of Japan, (in Japanese).
- 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.
- Cairns, C. E., & Feinstein, M. H. (1982). Markedness and the theory of syllable structure. *Linguistic Inquiry*, *13*, 193-225.
- 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.
- Connine, C. M., Blasko, D. G., & Wang, J. (1994). Vertical similarity in spoken word recognition: Multiple lexical activation, individual differences, and the role of sentence context. *Perception & Psychophysics*, *56*, 624-636.
- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 113-121.

- Cutler, A., & Otake, T. (1994). Mora or Phoneme? Further evidence for language-specific listening. *Journal of Memory and Language*, *33*, 824-844.
- 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. (1996). Computational models of spoken word recognition. In T. Dijkstra & K. de Smedt (Eds.), *Computational psycholinguistics* (pp. 114-138). London: Taylor & Francis.
- 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.
- Fudge, E. C. (1969). Syllables. *Journal of Linguistics*, *5*, 253-286.
- 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.
- Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, *28*, 267-283.
- Halle, M., & Vergnaud, J. (1980). Three dimensional phonology. *Journal of Linguistic Research*, *1*, 83-105.
- Hamburger, M. B., & Slowiaczek, L. M. (1996). Phonological priming reflects lexical competition. *Psychonomic Bulletin and Review*, *3*, 520-525.
- Kirtley, C., Bryant, P., MacLean, M., & Bradley, L. (1989). Rhyme, rime, and the onset of reading. *Journal of Experimental Child Psychology*, *48*, 224-245.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence: Brown University Press.
- Lenel, J. C., & Cantor, J. H. (1981). Rhyme recognition and phonemic perception in young children. *Journal of Psycholinguistic Research*, *10*, 57-67.
- 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.
- MacKay, D. G. (1972). The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology*, *3*, 210-227.
- Marslen-Wilson, W. (1984). Function and process in spoken word recognition---A tutorial review. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X* (pp. 125-150). London: LEA.

- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, **25**, 71-102.
- Marslen-Wilson, W. D. (1989). Access and integration: Projecting sound onto meaning. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 3-24). Cambridge: MIT Press.
- Marslen-Wilson, W. D. (1990). Activation, competition, and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 148-172). Cambridge: MIT Press.
- Marslen-Wilson, W. D., Moss, H. E., & van Halen, S. (1996). Perceptual distance and competition in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, **22**, 1376-1392.
- 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.
- Mehler, J., Dommergues, J. Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, **20**, 298-305.
- 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.
- Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language*, **32**, 258-278.
- 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.
- Sebastian-Galles, N., Dupoux, E., Segui, J., & Mehler, J. (1992). Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory and Language*, **31**, 18-32.

- Shillcock, R. (1990). Lexical hypotheses in continuous speech. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 24-49). Cambridge: MIT Press.
- Slowiaczek, L. M., & Hamburger, M. B. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1239-1250.
- 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.
- Stemberger, J. P., & Treiman, R. (1986). The internal structure of word-initial consonant clusters. *Journal of Memory and Language*, *25*, 163-180.
- Torretta, G. M. (1995). The "Easy-Hard" word multi-talker speech database: An initial report. *Research on Spoken Language Processing, Progress Report 20*, pp. 321-334. Bloomington, IN: Indiana University, Speech Research Laboratory.
- Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition*, *15*, 49-74.
- Treiman, R. (1985). Onsets and rimes as units of spoken syllables: Evidence from children. *Journal of Experimental Child Psychology*, *39*, 161-181.
- Treiman, R. (1986). The division between onsets and rimes in English syllables. *Journal of Memory and Language*, *25*, 476-491.
- Treiman, R., & Danis, C. (1988). Short-term memory errors for spoken syllables are affected by the linguistic structure of the syllables. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 145-152.
- Treiman, R., Fowler, C. A., Gross, J., Berch, D., & Weatherston, S. (1995). Syllable structure or word structure? Evidence for onset and rime units with disyllable and trisyllable stimuli. *Journal of Memory and Language*, *34*, 132-155.
- Treiman, R., & Zukowski, A. (1996). Children's sensitivity to syllables, onsets, rimes, and phonemes. *Journal of Experimental Child Psychology*, *61*, 193-215.
- Vroomen, J., & de Gelder, B. (1995). Metrical segmentation and lexical inhibition in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 98-108.
- Watanabe, T. (1996). The reaction time for the discrimination of nontarget words. *The Journal of the Acoustical Society of Japan (E)*, *17*, 323-324.
- Zwitserslood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition*, *32*, 25-64.

Appendix

balm	ban	bead	beak	bean	both	bud	bug	bum	bun	cause	chain	chat
check	cheer	chief	chore	cod	comb	con	cot	curve	dame	death	deep	den
dirt	does	dog	doom	down	dune	fade	faith	fig	fin	firm	five	food
fool	full	gas	gave	girl	give	goat	god	gut	hack	hag	hash	hick
hid	hoot	hum	hung	hurl	jack	job	join	judge	kin	king	kit	knob
lace	lad	lame	league	learn	leg	lice	live	long	lose	love	mace	main
mall	mat	mid	mitt	moan	moat	mole	mouth	move	mum	neck	noise	pad
page	pat	path	pawn	peace	pet	pool	pull	pup	put	rat	reach	real
rhyme	rim	roof	rough	rout	rum	rut	sane	serve	shall	shape	ship	shop
sill	size	soak	soil	south	suck	tan	teat	teeth	theme	thick	thing	thought
toot	vice	voice	vote	wad	wade	wail	was	wash	watch	wed	weed	whore
wick	wife	work	young									