Encoding operation and transcoding as the major loci of the frequency effect

In-mao Liu,*, Jei-tun Wu, Tai-li Chou

National Chung-Cheng University, Graduate Institute of Psychology, Ming-Hsiung, Chia-Yi, 621, Taiwan
National Taiwan University, Taipei, Taiwan

Received 10 January 1994, final version accepted 16 August 1995

Abstract

The present study used a lexical naming task as well as a regular naming task and a lexical decision task for locating the frequency effects in lexical decision and naming. The naming of Chinese characters in the lexical naming task (pseudocharacters also presented as in the lexical decision task) involves decision processes, while they are absent in the regular naming task. Since naming a Chinese character necessarily involves lexical access, a decision component of the frequency effect in lexical decision can be isolated. This procedure will not work for alphabetic orthographies, because sublexical processes underestimate the frequency effect in the regular naming task. As a consequence, not only can a prelexical component of the frequency effect in lexical decision be estimated, but a postlexical component of the frequency effect in naming can be estimated.

1. Introduction

In reading a written passage, one of the most fundamental and important processes is accessing the stored representation of words known as lexical access. Waley (1978) investigated a large sample of words with the lexical decision task with the aim of determining which properties of words had the strongest influence on speed of response, and came to the conclusion that frequency of occurrence was by far the most important variable. Evidence that subjects respond faster to high-frequency words than low-frequency words has accumulated (e.g., Becker, 1976; Berry, 1971; Forster & Cham-

The question of where the word frequency effect is located is, however, still controversial. An early theoretical view is that the frequency effect is localized in the way a word is stored or searched. Thus, Morton (1969, 1970) assumed that low-frequency words have higher thresholds for word recognition than high-frequency words. According to Becker (1979, 1980), Forster (1976), and Glanzer and Ehrenreich (1979), low-frequency words are searched only after high-frequency words have been searched.

1.1. Frequency effects obtained in different tasks

Subsequent studies (e.g., Balota & Chumbley, 1984, 1985; Chumbley & Balota, 1984; Forster & Chambers, 1973; Frederiksen & Kroll, 1976) showed that the size of the frequency effect is much larger in a lexical decision task than in a naming task. In considering the locus of the frequency effect, then, it becomes important to ask which task yields a dependable frequency effect, because either the lexical decision task will overestimate the true frequency effect or the naming task will underestimate it.

A general view is that the frequency effect in the naming task reflects the true influence of frequency on access time, and the larger effect in lexical decision reflects an additional decision component (e.g., Balota & Chumbley, 1984, 1990; Chumbley & Balota, 1984).

On the other hand, according to two-process theories (e.g., Coltheart, 1978; Henderson, 1982; Meyer, Schvaneveldt, & Ruddy, 1974; Seidenberg, 1985a), a word may be named by two parallel processes whose output is continuously pooled: (a) retrieval of a learned pronunciation following a lexical identification, and (b) assembly of pronunciation from knowledge of sublexical spelling-to-sound correspondences. Monsell, Doyle, and Haggard (1989, p. 49) reasoned that “the more the correct pronunciation of a word must depend on unique identification of the word rather than general knowledge of spelling–sound correspondences, the closer the effect of frequency observed for naming latency approximates that observed for lexical decision”. These investigators apparently believed that the lexical decision task gives the true frequency effect, and that the pronunciation task usually underestimates.

In support of this two-process argument, Frost, Katz, and Bentin (1987) observed that naming latency showed a smaller frequency effect for readers of Serbo-Croatian (a shallow orthography) than for readers of English or Hebrew (a correct pronunciation requiring its identification in a deep orthography), whereas lexical decision times did not show the same pattern. However, the frequency effect obtained in naming was still smaller than that obtained in lexical decision across the three orthographies: 30 versus 80 ms
for Serbo-Croatian, 50 versus 150 ms for English, and 40 versus 100 ms for Hebrew (measures estimated from Figure 1 of their article).

This means that for all alphabetic orthographies a lexical decision time difference between high- and low-frequency words is about two to three times that obtained for naming latency. This is perhaps the reason why it is generally understood that either the lexical decision task overestimates the frequency effect or the naming task underestimates it.

Let us then consider the relative size of the frequency effects in naming and lexical decision in Chinese orthography. Although a particularly large frequency effect in naming Chinese characters was early reported (142 ms for characters with pronunciation cues, phonograms; 248 ms for non-phonograms) by Seidenberg (1985b), its significance has not been noted for many years. The size of the frequency effect in naming Chinese characters is several times larger than that of the frequency effect (20–50 ms) in naming alphabetic words. Perhaps, most people expected that the frequency effect in lexical decision for Chinese characters would be even larger. More recently, Hue (1992) also obtained a frequency effect of 217 ms in naming characters with pronunciation cues (regular characters) and of 388 ms in naming characters with the conflicting pronunciation cues (irregular). Again, the lexical decision task was not included in his study. A study that includes both naming and lexical decision tasks in the same experiment is needed. Although Paap and Noel (1991) briefly mentioned that Paap, Chen, and Noel (1987) had used both lexical decision and naming tasks, the experimental details are not available.

1.2. Purpose of the experiment

1.2.1. General features of Chinese characters

Before describing the purpose of the present experiment, it is useful to know some general and specific features of Chinese characters. Each Chinese character represents one morpheme, which is generally one spoken syllable. The majority of Chinese morphemes are free in the sense that each can be used by itself as a word. Although Huang and Liu (1978) estimated more than 60% of words to consist of two characters in modern Chinese, out of the most frequently used 500 words obtained from the Liu, Chuang, and Wang count (1975) single-character words constitute more than 60%. One complication is that the same characters may be used as single-character words and also combined with other characters to form two-character words. This is why both the mean character frequency and the mean word frequency were calculated for a set of characters in the sequel.

Let us then consider the problem of how a Chinese character is pronounced. More than 80% of Chinese characters are phonetic compounds. These phonetic compounds are made up of a phonetic component and a semantic radical. In the examples shown in Table 1, each character consists of two groups of strokes: one group occupying the left half; and another
occupying the right half. The left part of each character is a radical, which contributes some aspect of its meaning to the character. The right part of each character is a phonetic component (or simply phonetic), which cues the sound of the character.

Depending on the phonetic cueing function of phonetics, all phonetic compounds may be classified into three types. For one type of character, its phonetic component is itself a character and is therefore pronounceable. In addition, its pronunciation is identical to that for the compound. In order to sharpen the present classification, this type of character is restricted to a case in which the intonation of its phonetic component is also identical to that of the compound. This is important, because different intonations give rise to different meanings in Chinese. This type of character will be called a “compatible phonetic”. For another type of character, its phonetic component is also a character but its pronunciation is different from that of the compound. A “different” pronunciation is restricted to a case in which both initial consonants and rhymes are different. This type of character is analogous to an exception word, and will be called an “incompatible phonetic”. For still another type of character, although it also consists of a radical and another component (stem), the latter component is not a character and therefore not pronounceable. The third type of character will be referred to as a “non-phonetic”.

Using English compound words as examples to represent radical–phonetic characters, if water-gate is a compatible phonetic character it is pronounced as gate (water is comparable to a radical), if water-pipe is an incompatible phonetic character it may be pronounced as craft, and if water-yxc is a non-phonetic character (yxc is neither a word in a dictionary nor pronounceable) it is pronounced as fowl. Examples of the three types of characters used in the present study are presented in Table 1.

Table 1
A sample set of various types of characters

<table>
<thead>
<tr>
<th>Character type</th>
<th>Freq.</th>
<th>Character</th>
<th>Pron.</th>
<th>PC Pron.</th>
<th>English meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compatible phonetic</td>
<td>High</td>
<td>評</td>
<td>ping (2)</td>
<td>ping (2)</td>
<td>Criticize</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>町</td>
<td>ding (1)</td>
<td>ding (1)</td>
<td>Watch</td>
</tr>
<tr>
<td>Incompatible phonetic</td>
<td>High</td>
<td>冰</td>
<td>bing (1)</td>
<td>suey (3)</td>
<td>Ice</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>恕</td>
<td>tien (2)</td>
<td>so (2)</td>
<td>Quiet</td>
</tr>
<tr>
<td>Non-phonetic</td>
<td>High</td>
<td>服</td>
<td>fu (2)</td>
<td>–</td>
<td>Dress</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>札</td>
<td>tsa (2)</td>
<td>–</td>
<td>Letter</td>
</tr>
</tbody>
</table>

Note: “English meaning” refers to the meaning of the character in English. Freq. = frequency, Pron. = pronunciation, PC Pron. = pronunciation of the phonetic component. A numeral in parentheses refers to intonations. 1 = flat, 2 = rising, 3 = falling then rising.
1.2.2. Characteristic character-to-sound processes

There is one characteristic feature of Chinese characters that may affect the relative size of the frequency effects in naming and lexical decision. In alphabetic orthographies, there is only a very limited alphabet, for instance, there are only 26 letters for representing about 44 phonemes. This means that the number of letters serving as cues for sounds is very limited in alphabetic orthographies, although grapheme–sound correspondences are not always one-to-one.

Now, let us consider Chinese characters formed by compounding sound-cueing phonetics and meaning-conveying radicals. There are 214 radicals and more than 800 phonetics. This means that the number of distinct script units or components serving as the cues for pronunciation is much larger for Chinese characters (over 800) than for alphabetic orthographies (26 for English). This fact has the following implications. For a low-frequency Chinese character, the subject has to select a correct phonetic component from among hundreds of sound-cueing components in order to pronounce it (carrying out a correct component-sound mapping) correctly. Consequently, the naming latency should be much slower for low-frequency Chinese characters than for low-frequency of alphabetic words. For a high-frequency item, it is well established that the subject is not likely to rely on the rule system for its pronunciation. Therefore, it is expected that its naming latency may not vary markedly across alphabetic and non-alphabetic orthographies.

The next problem is to consider how a character is pronounced through the aid of its phonetic. Although many radicals are pronounceable, they do not contribute any aspect of their pronunciation to the characters. It should be noted that the cueing function of phonetics for the pronunciation of characters is not rule-governed. Therefore, the relationship between phonetics and sounds in actual characters is not unambiguous (Zhou, 1978).

Since there are about 200 radicals and 800 phonetics in forming phonetic compounds, one may expect that more than 160,000 phonetic compounds can be generated. The number of characters in daily use is, however, only about 4500. The authoritative K’ang-hsi Dictionary of the eighteenth century lists between 40,000 and 48,000 characters, depending on the edition. Therefore, when a radical is arbitrarily combined with a phonetic, most often a “pseudocharacter” is created. A radical–phonetic combination is a pseudocharacter, simply because it is not an entry in a dictionary. If a radical–phonetic combination is a pseudocharacter, it has no pronunciation. This is because the pronunciation of a character is not governed by rules but must be learned by rote memorization. This is distinctly different from the case of representing sounds by letters of the alphabet.

In accordance with the foregoing analysis, Perfetti and Zhang (1991) demonstrated that there is no prelexical phonology in Chinese. Only entries in a dictionary have pronunciation in Chinese. An implication is that two-process theories are inapplicable to Chinese orthography: a phonologi-
1.2.3. Goal of the present experiment

The term *lexical access* may be used to refer to a complete processing sequence whereby a person gains access from an orthographic encoding of a word to stored properties such as its meaning and pronunciation. In the context of studying frequency effects, the term *lexical access* is generally used to refer to what Monsell et al. (1989) called *lexical identification*. This term refers to the process of contacting a stored word form without necessarily contacting other words properties. We use the term *transcoding* to refer to the process of gaining access from an orthographic encoding of a word to stored properties of its pronunciation.

We have noted that a radical–phonetic combination cannot be pronounced without recognizing it as a character. This means that an observed correct pronunciation of a character presupposes its lexical identification. If a character's pronunciation always follows its lexical access, a decision component can be isolated as follows.

First, devise a *lexical naming task* which is an analogue of the lexical decision task. The lexical naming task differs from a regular naming task in that pseudocharacters are included as well as characters. When an item (a character or a pseudocharacter) is presented in the lexical naming task, the subject is to name it if the item is a character. The subject is to say aloud "Chia" (meaning a pseudocharacter) if the presented item is a pseudocharacter, because it has no pronunciation.

Second, the naming of a character in the lexical naming task involves decision processes, while they are absent in the regular naming task. The single-route view guarantees that the naming of a character follows character recognition in both lexical and regular naming tasks. Therefore, how much of the frequency effect in the lexical naming task is due to a pure decision component can be measured by comparing the size of the frequency effects in the two naming tasks. This procedure will not work in English, because sublexical processes underestimate the frequency effect in the regular naming task.

Third, a frequency effect is defined by subtracting the mean reaction time of responding to high-frequency items from the mean reaction time of responding to low-frequency items. This means that part of the reaction time attributable to motor processes is eliminated in an observed frequency effect. Therefore, it is reasonable to assume that part of the frequency effect attributable to a decision component in the lexical naming task is identical to part of the frequency effect attributable to a decision component in the lexical decision task.

Fourth, since it is possible to measure how much of the frequency effect in
the lexical decision task is due to a pure decision component, a prelexical component of the frequency effect in lexical decision can be calculated. Furthermore, a postlexical component of the frequency effect in naming can similarly be identified. Thus, each component of the frequency effects in naming and lexical decision can be clearly specified.

The foregoing analysis is based on the characteristic feature of character-to-sound processes that the naming of a character presupposes its lexical identification. Another characteristic feature of character-to-sound processes is that there are more than 800 sound-cuing phonetics. It is predicted that the size of the frequency effect in lexical naming should be larger than that of the frequency effect in lexical decision.

2. Method

In those studies using English, it is well established that there is a Frequency × Regularity interaction in naming; that is, regularity effects (faster latencies for regular words compared to exceptions) are larger in low-frequency words and are small or non-existent in high-frequency words, although such interaction is not generally found in lexical decision (e.g., Andrews, 1982; Seidenberg, 1985b; Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Taraban & McClelland, 1987; Waters & Seidenberg, 1985).

The concept of regularity has to be defined somewhat differently for Chinese characters. As was explained earlier, the stimulus materials used in the present experiment were classified into three types: (a) compatible phonetic; (b) incompatible phonetic; and (c) non-phonetic. It is expected that a similar Frequency × Regularity interaction would be obtained in the naming tasks of the present experiment as in Hue’s (1992) and Seidenberg’s (1985b) studies. It is also expected that a Frequency × Regularity interaction would not be obtained in lexical decision as in previous English studies.

2.1. Subjects and materials

The subjects were 66 freshmen enrolled in introductory psychology courses at National Taiwan University. They participated in return for course credit.

Ninety high-frequency characters and 90 low-frequency characters were selected from the Wu and Liu (1987) count with the following restrictions. The high-frequency characters had a count of above 100, while a low-frequency character had a count below 15. Both high- and low-frequency characters were divided into three types: compatible phonetic, incompatible phonetic, and non-phonetic. For a character of the compatible phonetic type, its pronunciation is identical to its phonetic component in both sound and intonation. A character of the incompatible phonetic type is pro-
nounced differently from its phonetic component in both initial consonant and rhyme. For a character of the non-phonetic type, a component other than the radical is simply not a character and therefore not pronounceable.

In terms of these three types of characters, the mean frequency counts were 352, 414 and 348 (high frequency) and 8, 8, and 7 (low frequency) for the compatible phonetic, incompatible phonetic, and non-phonetic, respectively. Since these three types of characters were considered as a single-character words, their mean frequency counts as words were also obtained from Liu et al. (1975). They were 176, 155 and 185 (high frequency) and 1, 1, and 2 (low frequency) for the three types, respectively. It is to be noted that the latter count was based on a corpus of a million words and the former (Wu & Liu, 1987) based on a corpus of half a million words.

The three types of characters for both kinds of frequencies were equated with respect to initial sound and approximately equated for number of strokes. A total of 180 pseudocharacters was generated from the 180 characters by re-pairing radicals with phonetics in such a way that none of them was a character in a dictionary.

2.2. Tasks

There were three tasks: lexical decision task, regular naming task, and lexical naming task. The stimuli were presented, and subjects’ responses recorded, by an IBM-compatible computer. In the lexical decision task, the 180 characters and 180 pseudocharacters were intermixed and randomly presented one after another. The subject was to respond to each item by pressing a “True” key when it was a character and a “False” key when it was a pseudocharacter.

In the regular naming task, only the characters were presented. The subject responded to each character by naming it aloud into a microphone interfacing with the computer.

For the lexical naming task, both characters and pseudocharacters were used, as in the lexical decision task. The subject responded by naming the presented item when it was a character, or saying “Chia”, which meant in Chinese “pseudocharacter”, when the presented item was a pseudocharacter.

2.3. Design

Excluding those pseudocharacters, the design was essentially a $2 \times 3 \times 3$ factorial. The first factor was frequency of character, either high or low; the second was type of character, compatible phonetic, incompatible phonetic, or non-phonetic; and the third was kind of task, lexical decision task, regular naming task, or lexical naming task. The first two factors were with-subjects variables, while the last factor was a between-subjects variable.
2.4. Procedure

The subjects were randomly divided into three groups of 22 subjects each, according to order of appearance in the laboratory. Each group received one of the three tasks: lexical decision, regular naming, or lexical naming. The subjects were tested individually.

Each task started with 20 practice trials. The characters and pseudo-characters used on practice trials differed from those used on experimental trials. A trial began with an asterisk used as a fixation point in the middle of a computer screen for 250 ms. There was a 250 ms delay between the offset of the fixation point and the onset of the target item. The target remained on the screen until the subject responded by pressing a “True” or “False” key in the lexical decision task, or by naming it in the regular naming task, or by naming it or saying “Chia” in the lexical naming task. The subject had to say “Chia” in the lexical naming task, because a pseudocharacter is not pronounceable. During the practice trials, the feedback of “correct” or “wrong” was displayed for 800 ms on each trial. No such feedback was given on the experimental trials.

3. Results

Any observations deviating from the mean for each subject by more than two standard deviation units were excluded from the analyses, as were reaction times on the trials in which the items were misclassified or mispronounced.

The mean reaction times and error rates obtained under various conditions are presented separately for each task in Table 2. The mean reaction times and mean percentages of errors for pseudocharacters were 612 and 706 ms and 16.5% and 5.0% in lexical decision and lexical naming, respectively.

Several observations are apparent from the table. First, a clear Frequency × Regularity interaction is present in both regular naming and lexical naming, but negligible or absent in lexical decision. Second, the size of the frequency effect in lexical naming is much larger than that of the frequency effect in regular naming (a 44 ms difference), indicating that part of the frequency effect attributable to a pure decision component is substantial. When the mean latency of responding to high-frequency characters in regular naming is simply compared with that of responding to the same high-frequency characters in lexical naming, there is a large increase of 107 ms, because of the time involved in making decisions, even in naming high-frequency characters. Third, the frequency effect in lexical naming is much larger than that in lexical decision, reflecting the fact that the number of sound-cueing components (more than 800 in Chinese vs. 26 in English) could produce its effect on the latency of naming low-frequency items.
Table 2
Reaction times in milliseconds as a function of task, character type, and frequency

<table>
<thead>
<tr>
<th>Kind of task</th>
<th>Frequency</th>
<th>Type of character</th>
<th>C-P</th>
<th>I-P</th>
<th>N-P</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexical decision</td>
<td>High</td>
<td></td>
<td>446</td>
<td>455</td>
<td>454</td>
<td>452</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(2.0)</td>
<td>(3.9)</td>
<td>(1.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td></td>
<td>544</td>
<td>555</td>
<td>570</td>
<td>556</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(10.7)</td>
<td>(14.3)</td>
<td>(12.4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Diff.)</td>
<td></td>
<td>98</td>
<td>100</td>
<td>116</td>
<td>105</td>
</tr>
<tr>
<td>Regular naming</td>
<td>High</td>
<td></td>
<td>450</td>
<td>451</td>
<td>443</td>
<td>448</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1.6)</td>
<td>(3.1)</td>
<td>(1.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td></td>
<td>538</td>
<td>678</td>
<td>639</td>
<td>618</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(5.4)</td>
<td>(21.1)</td>
<td>(16.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Diff.)</td>
<td></td>
<td>88</td>
<td>227</td>
<td>196</td>
<td>170</td>
</tr>
<tr>
<td>Lexical naming</td>
<td>High</td>
<td></td>
<td>555</td>
<td>565</td>
<td>546</td>
<td>555</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1.9)</td>
<td>(4.0)</td>
<td>(2.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td></td>
<td>703</td>
<td>808</td>
<td>797</td>
<td>769</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(8.3)</td>
<td>(25.1)</td>
<td>(21.6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Diff.)</td>
<td></td>
<td>148</td>
<td>243</td>
<td>251</td>
<td>214</td>
</tr>
</tbody>
</table>

Note: Percentages of errors are given in parentheses. C-P = compatible phonetic, I-P = incompatible phonetic, N-P = non-phonetic.

Two analyses of variance were performed separately on the mean reaction times and mean percentages of errors obtained for each task: one based on subject means (F1), and another on item means (F2). With respect to the mean reaction times obtained for the lexical decision task, subjects responded faster to high-frequency than to low-frequency characters in both the subjects and items analyses, F1(1, 21) = 135.81, p < .01; F2(1, 174) = 230.84, p < .01. Type of character was marginally significant by the subjects analysis, F1(2, 42) = 3.40, p < .05, but not significant by the items analysis, F2(2, 174) = 1.48. The interaction between frequency and type of character was not significant: F1(2, 42) = 1.04 and F2 < 1. As for the mean percentages of errors, subjects made more errors in responding to low-frequency than to high-frequency characters by both the subjects and items analyses: F1(1, 21) = 58.89, p < .01; F2(1, 174) = 46.23, p < .01. Type of character was significant by the subjects analysis, F1(2, 42) = 3.94, p < .05, but not significant by the items analysis, F2(2, 174) = 1.33. The Frequency x Type of character interaction was significant, F1 and F2 < 1.

With respect to the mean reaction times obtained in the regular naming task, as is clear from Table 2, frequency and type of character and their interaction were all significant by both analyses: F1(1, 21) = 134.69, p < .01, F2(1, 174) = 246.72, p < .01; F1(2, 42) = 33.98, p < .01, F2(2, 174) = 16.10, p < .01; F1(2, 42) = 30.85, p < .01, F2(2, 174) = 14.54, p < .01. As for the mean percentages of errors, frequency and type of character and their interaction were also significant by both analyses: F1(1, 21) = 151.00, p <
Finally, with respect to the mean reaction times obtained for the lexical naming task, the same pattern of statistical results was obtained here as in the naming task. Thus, frequency and type of character and their interaction were all significant by both analyses: $F_1(1, 21) = 173.34, p < .01, F_1(1, 174) = 181.51, p < .01; F_2(1, 42) = 9.25, p < .01, F_2(2, 174) = 6.17, p < .01; F_1(2, 42) = 5.03, p < .05, F_2(2, 174) = 5.41, p < .01$. With regard to the mean percentages of errors, the same pattern of statistical results was also obtained as for the naming task. Frequency and type of character and their interaction were all significant by both analyses: $F_1(1, 21) = 83.09, p < .01, F_2(1, 174) = 60.47, p < .01; F_2(1, 42) = 25.88, p < .01, F_2(2, 174) = 7.70, p < .01; F_1(2, 42) = 25.44, p < .01, F_2(2, 174) = 5.16, p < .01$.

4. Discussion

The results generally replicate the presence of the Frequency × Regularity interaction in naming but its absence in lexical decision obtained with English studies (Coltheart, Besner, Jonasson, & Davelaar, 1979; Waters & Seidenberg, 1985). As Waters and Seidenberg pointed out, however, the Frequency × Regularity interaction was obtained in lexical decision in a situation where their subjects needed to base their decisions on the results of phonological analyses when “strange words” were included in the experiment. Since phonological analyses are generally made after finding that a presented item is a character, the Frequency × Regularity (compatibility) interaction would not always be obtained in lexical decision with Chinese characters. To illustrate the fact that single Chinese characters cannot be used to generate their pronunciation reliably, Paap et al. (1987) presented 50 characters of very low frequency to 12 skilled Chinese readers. When they were unable to recognize the character, they were asked to guess a pronunciation based upon whatever subcharacter cues they could find. Paap et al. found that only 13% of the pronunciations were correct.

4.1. Components of frequency effects

4.1.1. Estimation of the decision component

The size of the frequency effect in lexical decision was 105 ms (Table 2), which does not depart considerably from the norms obtained with alphabetic orthographies. The absolute lexical decision times of a 450–600 ms range are also fairly common in alphabetic orthographies.

The mean naming latency of high-frequency characters in the regular naming condition was 448 ms. When pseudocharacters were included in the lexical naming task, the mean naming latency of the same set of high-
frequency characters became 555 ms. This substantial increase of 107 ms represents a decision time involved. To calculate part of the frequency effect attributable to the decision process in lexical naming (decision effect), the mean frequency effect in regular naming was subtracted from that in lexical naming to yielding 44 ms (=214 – 170 ms).

4.1.2. Estimation of the prelexical component

The frequency effect in lexical decision was found to be 105 ms. This frequency effect consists of two components: (a) a decision component; and (b) a prelexical component. Therefore, the prelexical component of the frequency effect in lexical decision is estimated to be 61 ms (=105 – 44 ms). Our logic in isolating a decision component in the frequency effect in lexical decision is based on the assumption of identical decision processes that can be purely inserted into a naming task (Sternberg, 1969). One may suspect that subjects in the lexical naming task were simply waiting for a pronunciation to make itself available for a fixed amount of time (706 ms = mean naming latency of pseudocharacters) after which the “non-word” response was made. This suspicion may be dispelled, if we note the following. First, if no decision process is involved in naming a character in the lexical naming task, then it is difficult to explain why there was an increase of more than 100 ms in pronouncing a high-frequency character in comparison with the regular naming task. Thus, subjects were not simply waiting for a pronunciation to make itself available but making some decision. Second, the same suspicion applies to a lexical decision task, because in this lexical decision task subjects could be simply waiting for a word to be recognized and a “Yes” response made for a fixed interval (612 ms = mean latency of responding to pseudocharacters), after which a “No” response was made. This was not the case, because Balota and Chumbley (1984) among others showed that a decision process is necessarily involved in lexical decision.

4.1.3. Estimation of the postlexical component

Both naming and lexical decision necessarily involve lexical access in Chinese. Therefore, if the frequency effect in lexical naming is larger than that in lexical decision, then this difference must represent a postlexical component, because it is a case in which subjects recognized a low-frequency word but were unable to pronounce it for some time. When estimated from the results of Table 2, this difference is 109 ms (=214 – 105 ms). This is a mean additional time used for retrieving pronunciations of low-frequency characters after making lexical access.

As is clear from the “Diff” rows in Table 2, there is a large variation in the frequency effects in naming different types of characters. The frequency effects are particularly larger in naming incompatible phonetic and non-phonetic characters than in naming compatible phonetic characters. This implies that the postlexical component of the frequency effect in naming varies considerably depending on type of character. However, Perfetti,
Zhang, and Berent (1992) estimated from a sample of 300 phonetic compounds that the validity of phonetic components is an inverse function of frequency. In other words, there are more compatible phonetic characters than others in the low-frequency range. Therefore, as a whole, Chinese characters may not be read so slowly as Table 2 shows.

The ratio of the frequency effect obtained for "compatible" or "regular" characters to that obtained for "incompatible" or "irregular" characters was 2:3 in our pilot experiments conducted earlier, but decreased to 1:2 or smaller in the present experiment that was conducted later in the academic year. Although the subjects served in the present experiment only once, they had participated in other naming experiments and had thus become more experienced in the naming situation.

A small ratio means a small frequency effect for characters with pronunciation cues. This makes sense because in reading characters aloud very few subjects are consciously aware that the phonetic component of a character provides cues for its pronunciation. Through the pressure to name characters as fast as possible in the laboratory, subjects may come to pay special attention to subcharacter cues for fast naming. This is because the number of frequently used phonetic components is much smaller than the number of characters. Therefore, an experimental finding that there is no prelexical phonology in Chinese (Perfetti & Zhang, 1991) may apply to some characters more than others and to an early rather than late stage of practice. By just noting the ratios (2:3) of these frequency effects obtained in Hue's (1992) and Seidenberg's (1985b) experiments, we can tell that their subjects did not have much experience in fast naming under pressure before appearing in the laboratory.

4.2. A model of frequency effects

Although the subjective experience of accessing a word representation is indisputable in the lexical decision task, it is difficult to mark when such experience occurs in the naming of alphabetic words. Perhaps this is the main reason why parallel distributed processing (PDP) models that do not assume the existence of a lexicon are now attracting the attention of many investigators. These models generally assume interconnecting modules consisting of orthographic, phonological, and hidden units. Weights on connections between units are modified during training. According to a PDP model proposed by Seidenberg and McClelland (1989), lexical decision times are obtained first by computing orthographic error scores and then by transforming them into decision times through some function. In a PDP model proposed by McClelland and Rumelhart (1985), the pattern of activation is computed to obtain the probability of choosing one response over others. The reaction time is then measured by observing when the probability reaches a particular threshold activation value.

For Chinese orthography, however, the presence of word-level repre-
presentations is plausible, because without being recognized as a character a presented item cannot be pronounced in the naming task. Nevertheless, an idea not incompatible with those models that do not assume any lexicon will be developed for explaining the present results as follows.

An observed lexical decision time (LDT), which depends on low- and high-frequency characters, LDT(low) or LDT(high), may be considered to consist of the following components: the first is an encoding operation starting from character presentation to computation of a wordness value that depends on frequency, WN(low) or WN(high); the second is from computing a WN value to making the decision of selecting a response, DECISION(low) or DECISION(high); and the third is from making the decision to executing a response, OUTPUT, which should not depend on frequency. A wordness value is the familiarity value of an item as it is perceived as a word. This wordness value is similar to the familiarity/meaningfulness value of Balota and Chumbley (1984), the activation level of McClelland and Rumelhart (1985), and the orthographic and phonological error scores of Seidenberg and McClelland (1989), although in the latter the error scores are computed in one step. As a result we have:

\[ LDT(\text{low}) = WN(\text{low}) + \text{DECISION}(\text{low}) + \text{OUTPUT} \]

\[ LDT(\text{high}) = WN(\text{high}) + \text{DECISION}(\text{high}) + \text{OUTPUT} \]

Frequency effect in lexical decision

\[ \text{Frequency effect in lexical decision} = WN(\text{low}) - WN(\text{high}) + \text{DECISION}(\text{low}) - \text{DECISION}(\text{high}) = \text{Prelexical + Postlexical (LD)} \]

where the prelexical component refers to \([WN(\text{low}) - WN(\text{high})]\) that takes place before WN values are computed, and the postlexical(LD) to \([\text{DECISION}(\text{low}) - \text{DECISION}(\text{high})]\) that takes place after computation of WN values.

With respect to naming times, depending on whether they involve the decision component as in the lexical naming task or not as in the regular naming task, similarly we have:

Frequency effect in regular naming

\[ = [WN(\text{low}) - WN(\text{high})] + [\text{TRANS}(\text{low}) - \text{TRANS}(\text{high})] = \text{Prelexical + Postlexical (Ci)} \]

Frequency effect in lexical naming
= Prelexical + \[\text{DECISION(low)} - \text{DECISION(high)}\] + Postlexical(Ci) \\
= Prelexical + Postlexical(LD) + Postlexical(Ci) \quad (5)

Regular naming consists of the following processes: the first process of word presentation to computing a WN value is essentially identical to that in the case of LDT; the second process is from computing a WN value to retrieval of the pronunciation to articulatory motor commands, which depends on a type of character (C1 for compatible phonetic, C2 for incompatible phonetic, and C3 for non-phonetic) and is identified as the transcoding process (TRANS), since the two subprocesses involved are not separable in the present manipulation; and the third process is to execute motor commands, which is considered not to depend on word frequency and disappears in calculating the frequency effect in naming. In the case of lexical naming, a decision process is involved in naming and Eq. (5) is obtained.

The proposed additive model of components of the frequency effects in lexical decision and naming was tested by the method of least squares. The model accounts for 97.8% of variance, \(F(4, 4) = 44.91, p < .01\). The five parameters estimated by the method of least squares were 61 ms for prelexical, 44 ms for postlexical(LD), 35 ms for postlexical(C1), 152 ms for postlexical(C2), and 141 ms for postlexical(C3).

4.3. Relations to existing findings and views

Since Balota and Chumbley (1984, 1985) questioned the view that lexical identification is a major locus of frequency effects, no studies have been able to isolate a decision component or postidentification processes involved in lexical decision and naming. In view of the present results, Balota and Chumbley were correct in pointing out that a substantial portion of the frequency effect in lexical decision stands for a decision component. In fact, about 40% of the frequency effect is due to decision processes.

4.3.1. Frequency effects located postlexically

Researchers have long been concerned with whether frequency that affects pronunciation latency influences the encoding of a word, its production, or both. Balota and Chumbley (1985) conducted experiments in which a word was presented to the subject, and then, after a delay, the subject was cued to pronounce the word as quickly as possible. Since significant frequency effects were obtained at the long delays, they reasoned
that the observed frequency effect in delayed naming should be attributed, not to lexical access, but to a later process (production). However, this conclusion is indirect, because it assumes that there is always lexical access in word naming and that there cannot be lexical re-access in response to a cue to pronounce an item in delayed naming. Moreover, when the stimuli were equated in terms of articulatory properties, McRae, Jared, and Seidenberg (1990) found that large frequency effects obtained in immediate naming were eliminated at the longest delay intervals.

McCann and Besner (1987) found that pseudohomophones were named faster than other non-words, which suggests that a phonological lexicon must be contacted. However, the size of this advantage did not depend on the frequency of the words of which they were homophones. By reasoning that it is unparsimonious to suppose that one lexicon is frequency sensitive and the other not, they concluded that the only place left to locate the frequency effect is in the linkage between orthographic and phonological lexica. Thus, their evidence for locating the frequency effect in naming in a postlexical process is also indirect. In addition, Taft and Russell (1992) were able to obtain a frequency effect in the naming of pseudohomophones, when they controlled for orthographic factors that McCann and Besner ignored.

In comparison to the indirect evidence provided by Balota and Chumbley (1985) and McCann and Besner (1987), the results of the present experiment show that part of the frequency effect in naming Chinese characters may be localized in the linkage between orthographic and phonological lexica and/or the latter. This postlexical component was estimated by the additive model of Eqs. (3), (4), and (5) to be quite large, about 150 ms for incompatible and non-phonetic characters. This kind of reasoning will be difficult to make with alphabetic orthographies, because they allow a possibility for prelexical phonology.

4.3.2. Frequency effects located in lexical access

Balota and Chumbley (1984) argued that decision processes having little to do with lexical access exaggerate the word frequency effect in the lexical decision task. Their argument is based on their finding that the time to classify a word as belonging to a particular semantic category was not significantly affected by the frequency of that word. Although decision processes are undoubtedly involved in the lexical decision task, their evidence is quite indirect because it is well documented that the semantic categorization time depends largely on how many features are shared by the category and the exemplar, rather than on the frequency of the latter, especially when many categories are used in an experiment. For example, when the category and the exemplar share many features, categorization time is fast on positive trials but slow on negative trials. Thus, when the subjects were required to classify words as belonging to one of two very
large and general categories, either “human” or “inanimate”, Monsell et al. (1989) obtained a strong frequency effect.

With respect to Chinese orthography, part of the frequency effect that should be attributed as prelexical can be estimated rather clearly. The theoretical value of this prelexical component estimated from the model was the same as the value calculated from the observations, that is, about 60 ms.

After controlling for word length, in an eye movement monitoring study Rayner and Duffy (1986) obtained about a 40 ms frequency effect on first fixation duration and about a 80 ms frequency effect on gaze duration. If first fixation duration is taken as an index for lexical access time, then a 40 ms frequency effect is still substantial. This value could be an underestimate, because fixation duration is always measured in sentence contexts. Moreover, eye movement studies are problematic because it is difficult to differentiate between (or localize the effect solely to) lexical access or sentence integration processes.

The present estimation of lexical access time is, however, compatible with what was estimated from eye movement studies (e.g., Rayner & Pollatsek, 1989). Rayner and Pollatsek concluded from their review that a normal fixation time of 200–250 ms comprises about 50 ms for the stimulus to be registered in the brain, another 80–120 ms to encode the word, and (by subtraction) another 50–70 ms to decide that an eye movement should be made and begin its execution (p. 176). According to the present analysis, a normal lexical decision time of 450 ms for a high-frequency character requires about 50 ms for the character to be registered in the brain (this component should be the same), about 110 ms to decide which key to press, another 200 ms to execute a motor response, and still another 70–110 ms to encode the word (lexical access time, by subtraction).

5. Conclusion

The reaction time measures obtained in lexical decision and naming may be analysed to consist of the following components. First, word identification consists of the encoding operation which is the major locus of the frequency effect and is separable from the decision component in lexical decision. Second, the result of word identification is transcoded from orthography to phonology in naming, which accounts for an additional locus of the frequency effect in naming. Together with the decision component, these three loci of the frequency effect can be isolated, because naming in Chinese necessarily involves lexical access. The present study merely capitalizes on some distinctive surface features of Chinese characters to unravel components of the frequency effects in lexical decision and naming that have plagued investigators of alphabetic orthographies.
Acknowledgements

This study was supported by National Science Council Grant 83-0301-H-194-039. We thank three anonymous reviewers for many helpful comments on early versions of this paper.

References


In-mao Liu et al. / Cognition 59 (1996) 149–168


