Memory & Cognition
1991, 19 (3), 263-273

Depth of phonological recoding in short-term memory

YI XU
University of Connecticut, Storrs, Connecticut
and Haskins Laboratories, New Haven, Connecticut

In testing the hypothesis that surface phonetic form is included in short-term memory (STM) representation, the tone sandhi phenomenon in Mandarin Chinese was exploited, and, as a prerequisite, the hypothesis that tonal similarity affects STM of verbal material in a tone language was also tested. In Experiment 1, subjects recalled visually presented sequences of seven monosyllabic Chinese morphemes having either the same tone or different tones. More errors were made on the monotonal sequences than on the multitonal sequences, confirming the effect of tonal similarity on STM. In Experiment 2, subjects recalled visually presented sequences of disyllabic nonsense words. The sequences were designed in such a way that half of them were subject to the tone sandhi rule in Mandarin Chinese, whereas the other half were not. The consequence of applying the tone sandhi rule, as designed, was to make all the first characters in the sequences identical in pronunciation, thus creating potential phonological confusion. More errors, indeed, occurred on the sequences subject to the tone sandhi rule than on those not subject to it, indicating the existence of a surface phonetic representation in STM. The findings in this study provide further insight into the phonological mechanism of STM. Different accounts for this mechanism are also discussed in the light of the new findings.

It has been long established that linguistic material, whether presented auditorially or visually, has to be retained in speech form in memory if it is to be remembered for more than a few milliseconds. This has been evident from the results of various studies showing the phonological nature of short-term memory (STM). It has been found that the errors subjects make in immediate recall are mostly phonological confusions; that is, items that share the same phonological features with the stimulus items were used in erroneous responses instead of items that share visual or semantic properties with the stimulus items (Brady, Mann, & Schmidt, 1987; Conrad, 1964; Wickelgren, 1965a, 1965b). It has also been found that subjects make more errors recalling series containing phonologically similar items than they do recalling phonologically dissimilar series. This is true for both auditorially presented sequences and visually presented sequences (Baddeley, 1966, 1968; Conrad, 1964). In visual presentation, the effect of phonological similarity has been found to be independent of the nature of the stimuli, as long as they can be represented linguistically. It has occurred with written letters (Conrad, 1964), numbers or words (Baddeley, 1966, 1968), and nameable pictures (Conrad, 1972).

Neither does the effect of phonological similarity seem to vary with the nature of the writing system. Erickson, Mattingly, and Turvey (1977) found that Japanese kanji characters were harder for native speakers to recall correctly when the set of characters to be remembered had phonetically similar readings. Tzeng, Huang, and Wang (1977) showed that phonological similarity, whether in the visually presented items to be recalled or in the shadowing items that were auditorially presented and repeated by the subjects, interfered with STM for Chinese characters. Yik (1978) showed that the effect of phonological similarity on STM of Chinese characters is larger than the effect of visual similarity, although the latter was also found to be significant.[1] Zhang and Simon (1985) found that the STM span for Chinese radicals without common names was no greater than three items, and the STM span for complete homophones[2] was the same as that for radicals. But the STM span for nonhomophones was significantly greater than that for homophones. Ren and Mattingly (in press) investigated STM for Chinese characters in relation to the reading ability of second-grade primary-school children. They found that good readers were penalized more by phonological similarity in the series to be recalled than were poor readers. But visual similarity had no effect on either the good readers or the poor readers.

STM has also been shown to be sensitive to different aspects of phonology. Most of the studies showed rhyming effects upon STM (Baddeley, 1966; Erickson et al., 1977; Shankweiler, Liberman, Mark, Fowler, & Fisher, 1979). Consonantal similarity was also found to have a

This research was supported by NICHD Grant HD-01994 to Haskins Laboratories. I thank Ignatius G. Mattingly, Alvin M. Liberman, Bruno H. Repp, Arthur S. Abramson, Diane Lillo-Martin, and Robert G. Crowder for their valuable comments on an earlier version of this paper. Correspondence should be addressed to the author at the Department of Linguistics, University of Connecticut, Storrs, CT 06269.

significant effect on STM (Brady et al., 1987), although some studies found that the effect of consonantal similarity is much smaller than that of vowel similarity (Tzeng et al., 1977).

The phonological nature of STM implies, as shown by some of the studies mentioned above, that if the material to be remembered is orthographic, it has to be first transformed into a speech-like form. This kind of transformation process is usually referred to as phonological recoding. The term phonological recoding is, however, ambiguous in that the levels of phonology involved are not specified. It could be understood as the underlying morphophonemic level, at which the phonological representation of a linguistic unit (e.g., a morpheme or a word) is the one defined by the lexicon; it could be understood as the surface phonetic level, at which linguistic material is in a form similar to that generated in overt speech; or, it could be understood as including both. To determine which level is involved in STM, it is necessary to find cases in which there are different representations for the linguistic material to be remembered at the two levels. The representations could be different because of differences in the specifications that are unique to either level. It has been shown by Baddeley, Thomson, and Buchanan (1975) that when the number of syllables and the number of phonemes in words are held constant, STM span is inversely related to the temporal duration of the vowels in words, indicating a use of the surface specifications of speech as memory codes. The underlying and the surface forms could also be different because of discrepancies between the intended phonological representation and the actual surface sounds. Locke and Kutz (1975) showed that for children who mix certain contrasting sounds in their speech but correctly distinguish the same contrast in their perception, it is the surface sound they would produce, rather than their phonological intention, that provides codes for their short-term memory.

Discrepancies between the morphophonemic representation and the surface phonetic representation of linguistic material are more commonly seen when smaller linguistic units are combined to form larger ones; for instance, when morphemes are combined to form words and when words are combined to form phrases and sentences. This combination often results in a different number of distinctive categories at the two levels because of the application of phonological rules. This phenomenon can be employed to check whether surface phonetic representation or underlying morphophonemic representation is used as a code for STM. One way to do this is to find linguistic material that has more similarity in its surface forms than in its underlying forms and to employ the common finding that phonological similarity impairs the immediate recall of linguistic material. This is the approach used in the current research. This approach calls for the testing of linguistic material that contains more similarity in the surface phonetic representation than in the underlying morphophonemic representation. Specifically, the tone sandhi phenomenon in Mandarin Chinese was manipulated to create this material. Before presenting the details of the experiments, however, it is desirable to introduce some relevant background about the Chinese language family, as well as about Mandarin Chinese.

The Chinese languages are all tone languages. In a tone language, differences in the change of pitch provide not only intonational information about a sentence (e.g., whether it is a question or a statement) but also distinctions between lexical items. In other words, two utterances otherwise identical in pronunciation might represent completely different lexical items if they have different pitch patterns, or "tones." In Chinese languages, the basic units that carry tone patterns are the monosyllabic morphemes. Each monosyllabic morpheme is represented orthographically by a "zi," that is, a character, which is written separately from its neighboring characters (in the following discussion, "morpheme" and "character" are used interchangeably, depending on the context). Thus, the pronunciation of each character is defined not only in terms of consonants and vowels but also in terms of the tone associated with it. Structurally, the orthography of Chinese characters is often composed of two parts, each with a certain arrangement of strokes. One of them is often referred to as the "radical," which carries semantic information. The other part is often referred to as the "phonetic," which carries phonological information. However, probably because of the changes accumulated over the years, information about pronunciation carried by the phonetic is anything but accurate. Moreover, no information about the tone of a character is directly carried by the phonetic. That information is available only after the recognition of the morphemic identity of the character.

There are a number of different languages and dialects in the Chinese language family, and each of them has a different tonal system. The language used in this study is Mandarin, which is the official language both in mainland China and in Taiwan. In Mandarin, there are four tones,[3] which are conventionally called Tone 1, Tone 2, Tone 3, and Tone 4. In "pinyin," a quasiphonemic transcription system for Chinese characters, and also the transcription system used in this paper, the four tones are represented as marks over the nuclear vowels in the syllables. For example, *bā*, *bá*, *bǎ*, *bà* are the pinyin transcriptions of the syllable *ba* with Tones 1, 2, 3, and 4, respectively. Incidentally, those marks actually mimic the pitch contour of the four tones spoken in Mandarin.

In modern Chinese, the pronunciations of most of the morphemes are not unique. At the level of morphemes, homophones are commonplace. The number of different syllables with different tones recorded in *Cihai* (a grand dictionary of Chinese words; Xia, 1979) is 1,353, but the number of characters represented by those syllables in that dictionary is 14,872. Thus, an average syllable with tone represents about 11.02 different morphemes. To make things even worse, the information load of the syllables is far from being evenly distributed. For example, the syllable *néng* represents only one single morpheme, namely,

能

(to be able to, can). The syllable *yì*, on the other hand, represents 149 different morphemes, as recorded by *Cihai* (Xia, 1979). Fortunately, most of the Chinese words consist of more than one morpheme. Thus, the number of homophones at the word level is much smaller than that at the morphemic level. However, the very fact that the Chinese orthography is structured at the morphemic level rather than at the word level, and that there are a large number of homophones at the morphemic level, makes some of our investigations easier or even possible.

Mandarin has a phonological rule called the "Third Tone Sandhi" rule. The rule says that whenever two Tone 3s are in succession within the same word and/or the same rhythmic unit, the first tone must change into a tone that is indistinguishable from Tone 2. Wang and Li (1967) found that Chinese listeners could not distinguish between a Tone 2 derived from Tone 3 by the tone sandhi rule and a nonderived morphemic Tone 2. They played, to Chinese subjects, pairs of bimorphemic words that were identical phonologically, except that in one of them, the first morpheme is in Tone 2, whereas in the other, the first morpheme is in Tone 3, as in, for example, *máimǎ* (to bury a horse) and *mǎimǎ* (to buy a horse). The subjects' identification of those words was found to be very close to chance level.

The phonological ambiguity caused by application of the tone sandhi rule can be used in determining whether surface phonetic representation is generated in STM. If the surface representation is generated in STM, similarity at this surface level should adversely affect performance on immediate recall tasks. To verify this, certain linguistic material can be created in which the amount of tonal similarity varies between surface phonetic form and underlying morphophonemic form caused by the application of the sandhi rule. In other words, proper combinations of different tones can be manipulated so that there is more similarity in the material to be recalled after the application of the tone sandhi rule than there is before its application. If the surface phonetic representation is produced, then immediate recall for this kind of material should be poorer than for material to which no such tone sandhi rule is applicable. If the rule is not applied, that is, if the surface phonetic representation is never produced, immediate recall should be the same for both kinds of material.

Before the test for tone sandhi is carried out, however, it is desirable to make sure that tonal similarity affects STM at all, which to this date has not been confirmed experimentally. Hence, two experiments are performed in this study. The first one investigates the effect of tonal similarity on STM; the second one explores whether surface phonetic tonal similarity is produced through the application of the tone sandhi rule.

EXPERIMENT 1

Method

Material. Forty-eight test lists, each list consisting of seven Chinese characters and each character representing a monomorphic word, were used as stimuli in the experiment. In each list, all the characters shared the same vowel and final consonant but differed in initial consonants. In other words, the characters in each list rhymed with one another. Two rhymes, *an* and *u*, were used in the lists. Those two rhymes were used because there are a sufficient number of characters of high and moderate frequencies with all four Mandarin tones that have the two rhymes, so that enough morphemes with different initial consonants could be found for each of the lists. The test lists were of two types. Lists of the first type were composed of characters with the same tone; lists of the second type were composed of the same characters used in the lists of the first type but in different combinations, so that each of the lists consisted of rhyming characters with different tones. Because there are only four different lexical tones in Mandarin, some of the characters in the second type of list shared the same tone, but care was taken not to use any one tone more than twice in the same list.

Appendix A shows the sample lists that were used in Experiment 1. *List a* contains examples of the lists with the same tones, each of them consisting of characters with only one of the four lexical tones in Mandarin. The first four rows in *list a* all have the rhyme *an*; the last four rows all have the rhyme *u*. *List b* shows examples of the lists with different tones, each of them consisting of characters with different tones. As in *list a*, all the characters in the first four rows have the rhyme *an*, and those in the last four all have the rhyme *u*. The characters in *list b* are the same as those used in *list a* but in different combinations.

Subjects. In China, although everybody has a great amount of exposure to Mandarin[4] through radio and television, as well as varying amounts of training at school, most people speak their own language or dialect most of the time in their daily life until, in most cases, they enter college, where the common language is often Mandarin. Thus, it is not clear what kind of language or dialect they use when they read (assuming that they do use "inner language" of some sort in reading). As was mentioned in the introduction, each of the Chinese languages or dialects has its unique tonal system. Therefore, in a study involving tones, it is desirable not to mix subjects with different linguistic backgrounds, unless it is done deliberately.

Because the phonetic system of Mandarin is based on the native dialect spoken in Beijing, the speech of the people from Beijing, of all Chinese dialects, bears the closest resemblance to Mandarin. Thus, in any study on the pronunciation of Mandarin, native speakers of the Beijing dialect are the best subjects. Because this study was intended to test a hypothesis about the effect of phonological similarity on short-term memory, the ideal subjects were those who, not only in their conversation but also in their possible inner speech, speak the dialect whose phonetic system has the least deviation from the standard pronunciations that are recorded in the dictionary. Fortunately, there were a sufficient number of speakers of the Beijing dialect at the University of Connecticut for the purposes of this study. Twelve of them participated as subjects in Experiment 1. They were paid for their participation. Eight of them were females, and four were males. Most of them were graduate students at the University of Connecticut; the rest were college graduates. None of them reported any history of dyslexia.

Procedure. The method used in this experiment was similar to the paradigm introduced by Waugh and Norman (1965), which was used in many of the subsequent STM studies, including some of the studies on the STM of Chinese (Ren & Mattingly, in press). The characters in each of the 48 lists were displayed on the computer screen one at a time. Each character stayed on the screen for 1 sec, and the following character appeared on the screen 0.5 sec after the disappearance of the previous character. One and a half seconds after the last character in the list disappears, two short beeps were heard; then, another 1.5 sec later, a probe character appeared on the screen. This probe character was one of the characters that had occurred in the list. At this point, the subjects were required to write on the response sheet the character that followed the probe character in the list. Guessing was neither encouraged nor discouraged, and the subjects were allowed to leave a blank box on the answer sheet if they really could not remember the character. After each subject wrote down the character, the experimenter entered the pinyin transcription of that character into the computer. This input was later checked against the subjects' response sheet and corrected accordingly if there was any disagreement. Only one probe was used for each of the lists, and, because every list consisted of seven characters, there were six possible probe positions in each list. With 48 trials, four judgments were obtained for every position under each of the two tonal conditions. Both the order of the appearance of the lists and the position of the probe characters were selected randomly by the computer program and were different for different subjects.

The subjects were tested individually. Before the real test, each subject was given six practice trials to familiarize him or her with the procedure. At the very beginning of the test, each subject was given the following instructions in Chinese:

> This is a short-term memory test. During the test, Chinese characters will be shown on the screen one at a time. Please memorize every character, as well as its position relative to other characters in the same group of characters. At the end of each group of characters (seven), you will hear two short beeps. Then, a reference character will appear on the screen. This is one of the characters in the group just shown. At this point, please recall which of the characters in the group just shown had occurred after this reference character and write it down.

Note that in the instructions, no specific requirements were made as to whether the subjects should or should not pronounce the characters while trying to remember them. This was to allow the subjects to do the task in a way they felt was most natural to them. As it turned out, most of them did not pronounce the characters, although some of them did seem to be moving their lips during the test, as if trying to murmur to themselves.

Results

The subjects' responses with characters that were different from the stimulus characters but had pronunciations, including tones, identical to those of the stimulus characters were scored as correct responses. Figure 1 shows the percentages of correct responses against the serial order for the recall test. Unlike most STM studies, which typically show a serial-position effect, the pooled result here did not show an apparent serial effect. An ANOVA analysis showed that the main effect of serial positions was not significant and that there was no significant interaction between serial position and tonal similarity. Several factors might have contributed to this result. First, for each of the six serial positions in each of the two categories, there were only four trials. This was probably not enough to reveal a strong serial effect. Second, the beeps used before the presentation of the probe character could account for the absence of the normal rise in the last two positions. Some of the subjects reported that the beeps seemed to have wiped out their memory. Third, subjects seemed to have used different recall strategies when trying to remember the lists. Some of them reported that they had grouped the first four characters in the lists into one set and the last three characters into another set; others reported that they did the opposite; still others reported that they had tried to remember only the pronunciations of the first six characters, leaving the last one to visual memory. This could have resulted in different patterns of recall as a function of the serial positions of the items in the lists.

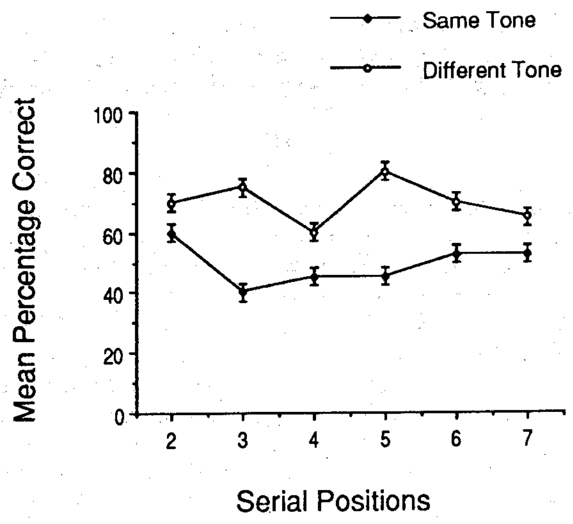

Figure 1. Mean percent correct responses for stimulus lists with monotonal and multitonal lists plotted against serial positions. (Horizontal bars represent standard error.)

The lack of a significant main effect of order did not affect the main effect of tonal similarity. As is obvious from Figure 1, the tonal similarity has a significant effect on the recall performance [$F(1,11) = 25.28$, $p < .001$]. When asked to comment on their experiences during the experiments, all of the subjects reported that the task was very hard, because the lists were very difficult to say, although only some of them had made apparent efforts to pronounce the lists quietly. Interestingly, none of the subjects realized that some of the lists were monotonal, whereas the others had different tones, and few of the subjects noticed that all of the lists rhymed internally. This is probably evidence that conscious phonological observation might not be directly related to the actual coding strategy used in reading.

Discussion

The results of Experiment 1 indicate that tone, like rhyme, plays an important role in STM for Chinese. This conclusion is not very surprising, because, in a tone language, the tonal features are part of the phonological specifications that enable one to distinguish between different morphemes. Baddeley (1966) found that vowel similarity in the word sequence to be recalled significantly affects the recall performance. Brady et al. (1987) reported that the number of errors in the ordered recall of auditorially presented syllables is a function of the phonetic similarity of consonantal features. The effect of tonal similarity on the STM of linguistic material in a tone language is simply a further demonstration that various aspects of the phonetic system, as long as it is relevant to speech, are likely to be involved in verbal short-term memory.

Because Experiment 1 confirmed the effect of tonal similarity on the STM of the speakers of a tone language, this finding could now be exploited to test the hypothesis that the Mandarin tone sandhi rule is applied in STM by checking if the effect of tonal similarity shows up again, this time as a result of a special design of the tests in Experiment 2.

EXPERIMENT 2

Method

Material. The test items used in this experiment were 48 lists, each consisting of 3 disyllabic nonsense words. Pilot tests had shown that lists with 4 disyllabic nonsense words were a bit too difficult for the subjects to recall. This agrees with the finding of Zhang and Simon (1985) that STM span for two-character Chinese words was 3.83, whereas for isolated characters it was 6.38. In their study, the 2-character words were real words, whereas here they are nonwords, which should reduce STM span even further. Three 2-character nonwords was found to be an appropriate size for this experiment.

Appendix B shows all the test lists used in this experiment. The lists were designed so that most of the errors would be confusions on the tones of the first characters in the nonsense words. Each element of a list consists of two characters that, although meaningful in themselves, do not form a meaningful bimorphemic word. For example, the characters in column 1 of the first row in Appendix B mean, respectively, "frequent" and "willing to." This combination is not a meaningful word in Chinese. In each list, the first characters in all three words have the same syllable structure (i.e., the same rhyme, as well as the same initial consonant) except for their tones, which are either Tone 2 or Tone 3. The graphic shapes of the characters have no similarity to one another. The second characters in each list all share the same tone but have completely different syllable structures, that is, they share neither the same initial consonant nor the same rhyme among themselves or with the first characters that precede them. Two tones are used for the second syllables in the lists, depending on the classification of the test lists. In the first type of list (Sandhi List, Appendix B), the second syllables all have Tone 3. This means that, in these lists, whenever the first character is in Tone 3, the possibility exists that the tone sandhi rule might apply as it always does in speech. If this happens, the first character should become indistinguishable phonetically from the other first character(s) in the same lists that have an underlying Tone 2. In this way, the three first characters in each list of this type would be phonetically identical to one another, resulting in possible confusion among them. In the second type of list (No-Sandhi List, Appendix B), the second characters all have Tone 4. In this case, no tone sandhi rule is applicable. Thus, the first characters in those lists will remain different in tone whether or not the tone sandhi rule applies in STM. As can be noticed, the first characters in the No-Sandhi Lists and the Sandhi Lists are exactly the same, but the second characters are all different. Thus, the only real difference between the two types of lists is in the tones of the second characters, which make the tone sandhi either applicable (in the Sandhi Lists) or nonapplicable (in the No-Sandhi Lists).

Subjects. Ten Chinese students, three male and seven female, participated in Experiment 2. Again, they were native speakers of the Beijing dialect. They were also paid for their participation.

Procedure. The procedure used in Experiment 2 was similar to that used in Experiment 1. Each of the disyllabic nonsense words was displayed on the computer screen for 1 sec. After a delay of .5 sec, the following word appeared on the screen. One and a half seconds after the disappearance of the last word in a list, two short beeps were heard and, 1 sec later, a probe word appeared on the screen and stayed there for 2 sec. The subjects' task was to write down the word that followed the probe word in the list just shown. When the probe item was • •, the subject was to write down the first word in the list. This was to guarantee that the subject would try to remember all three words in the list, rather than just the last two. This was necessary because, when there are only three words in the list, it is possible to ignore the first one and still accomplish the task. When the probe item is recognized as a strange word, the subject can safely assume it to be the first one in the list. This tactic was actually used, and rather successfully, by quite a few subjects in a pilot study, thus creating a confounding subcategory among the subjects.

The subjects were tested individually. Each subject was given six practice trials before the start of the real test. At the beginning of the experiment, written instructions in Chinese, similar to those used in Experiment 1, were given to the subject. As in Experiment 1, no requirements were made as to whether the subject should or should not pronounce the words, thus allowing the subject to do the task in a natural way. Some of the subjects did move their lips quietly during the test; others remained totally silent. After the subject wrote down each response, the experimenter entered it into the computer. This input was later checked against the subjects' response sheets and corrected accordingly. As in Experiment 1, guessing was not discouraged, but neither was it encouraged, and the subjects were allowed to leave a blank box on the answer sheet if they really did not think they could remember the word.

Results

The subjects' responses were scored in three different ways: First, any disagreement between the pronunciation of the response character and the stimulus character was counted as an error; second, only disagreements on the first characters were counted as errors; and third, only disagreements on tones of the first characters were counted as errors. The results of a two-factor within-subjects ANOVA for the data scored with those three methods were all very close. The following discussion will only use data obtained with the second method, that is, treating any disagreement on the first character as an error. Figure 2 shows the mean percent correct responses plotted against serial positions for the Sandhi Lists and the No-Sandhi Lists. Again, the effect of serial positions on STM was not significant. This time, the lack of a significant serial effect might be due to list length, which was not long enough to create clear contrasts among different serial positions. And, again, the beeps after the presentation of the word lists might have contributed to the absence of the final rising in the recall curve. The interaction between serial position and list type also was not significant.

There is a significant main effect of list type [$F(1,9) = 20.02$, $p < .01$]. The number of errors on the first

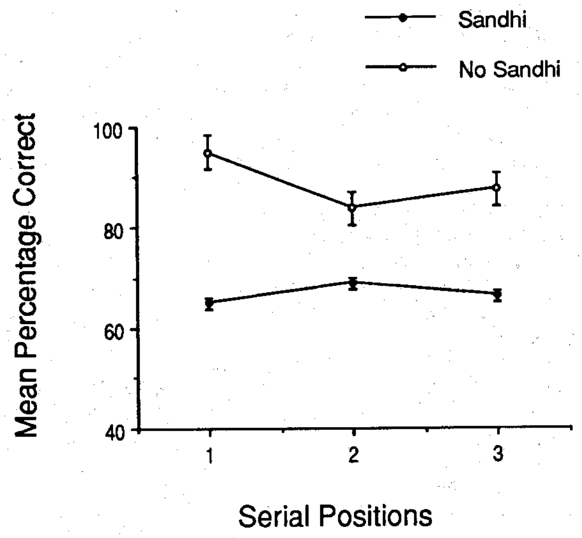

Figure 2. Mean percent correct responses plotted against serial positions for Sandhi Lists and No-Sandhi Lists. (Horizontal bars represent standard error.)

characters in the Sandhi Lists was two and a half times the number of errors on the first characters for the No-Sandhi Lists. Among all the errors on the Sandhi Lists, 77.2% were those on the tones of the first characters, whereas only 54.24% of the errors on the No-Sandhi Lists were on the tones of the first characters.

Almost all the errors on the tones of the first characters were confusions between Tone 2 and Tone 3. This is true for both the Sandhi and the No-Sandhi Lists. Of the total of 73 errors on the tones of the first characters in the Sandhi Lists, 43 were substitutions of Tone 2 for Tone 3 and 30 were substitutions of Tone 3 for Tone 2. A one-factor repeated measure ANOVA showed no significant difference between the two, indicating that the confusions were in both directions.

Discussion

Logically, to fulfill this particular immediate recall task, the best thing to remember would have been the visual forms or the semantic values of the characters, because they were all different from one another, so there was little possibility of confusion. The second best thing to remember would have been the underlying phonological forms, because there were at least two different tones on the first characters in the Sandhi Lists. But the subjects seemed to have used the worst thing, that is, the surface phonetic form, which was identical within each Sandhi List for the first characters. This can only imply that in performing the immediate recall task, the derivation of the surface phonetic forms through the application of the phonological rules is somehow compulsory.

When asked to comment on the experiment, most of the subjects reported that some of the lists seemed easier to remember because the first characters in those lists all seemed to be the same, whereas in some other lists, the first characters seemed to have different tones, so the subjects had to pay special attention to remember the tonal differences. Apparently, the three first syllables in each of the Sandhi Lists were taken by the subjects as having the same tone as well as the same segmental structure. Two things here are of interest to us. First, by reducing the difference between the pronunciations of the characters to only two tones, speakers of a language with a nonalphabetic writing system seemed to have been forced to come to awareness about some of the phonological features in their language. Second, the tone sandhi rule is so natural, even in a process that does not require overt phonetic production, that although all of the subjects applied it, they were nevertheless totally unaware of its occurrence.

GENERAL DISCUSSION

The two experiments performed in this study reveal a strong effect of tonal similarity on STM for visually presented material in a tone language. This effect is found to be present both when the surface phonetic form and the underlying morphophonemic form agree and when the two forms do not agree. In the latter case, the similarity in surface phonetic form affected STM in the absence of the same similarity in underlying morphophonemic form.

The results of this study further confirm the finding of previous studies that sensitivity of STM to phonological similarity is not a function of the nature of the writing system for the language being tested (Erickson et al., 1977; Hue & Erickson, 1988; Ren & Mattingly, in press; Tzeng et al., 1977; Yik, 1978; Zhang & Simon, 1985). The findings of this study also add a new aspect of the phonology, namely, tonal features in tone languages, to the list of the feature classes to which STM is sensitive. Furthermore, besides investigation of the effect of phonological similarity in general, the question about the levels of the phonology involved in STM is also addressed in this study. The results of Experiment 2 demonstrate that surface phonetic representation is generated in STM for visually presented material and that codes in surface phonetic form are actually used in STM.

There are two aspects of the findings in this study that are worth noticing. First, the tonal features of Chinese are, compared with other aspects of its phonology, even less directly represented in the orthography for this language, for even the phonetic radical of a character does not give any reliable clue about the tone of the morpheme represented by the character. Nevertheless, the tonal features of the characters were apparently activated during the tests, because the tonal similarity in the test material severely impaired the immediate recall of the material. Second, not only were the tones of the isolated characters activated, but also the tone sandhi rule was apparently applied, because the tonal similarity that is possible only after the tone sandhi also seriously impaired the recall of the material. Two questions could be asked. First, why should aspects of the phonology that have no direct representation in the orthography be activated at all? Second, why should the readers or recallers go all the way to generate the surface phonetic representation rather than stick to the underlying morphophonemic representation which, in that particular test situation, actually provides more relevant information for distinguishing between the test items? Evidently, in performing the STM task, the subjects were resorting to a phonological mechanism similar to the one used in speech production and speech perception. In other words, it seems that so-called phonological encoding involves not only the underlying abstract phonology of a language but also the surface phonetic representation which includes, though not necessarily in real acoustic forms, all the relevant features crucial to speech production and speech perception.

One account for the phonological nature of STM is the one proposed by Baddeley and Hitch (1974). They described the phonological mechanism as a buffer in the working memory. Controlled by the central executive component of the working memory, the phonological mechanism is said to be responsible for storing material for a few seconds while awaiting higher-level processing. A more recent attempt to explain the nature of the phonological mechanism in STM has been made by Ren and Mattingly (in press). They argued that the phonological mechanism is a consequence of the operation of the language module. According to their account, so-called short-term memory is simply a process for exploiting the language module, which is used both in the primary process of speaking and listening and in the secondary process of reading and writing. As shown by both current and previous research, in situations where visually presented linguistic material is to be remembered for more than a few milliseconds, the language module is also employed, which is perhaps the only choice, considering the very limited capacity of the short-term iconic storage for temporally presented information (Averbach & Coriell, 1961; Frick, 1985; Zhang & Simon, 1985). The nature of the language module is such that the computation and output of the various representations is automatic and compulsory (Fodor, 1983, Liberman & Mattingly, 1985; Mattingly & Liberman, 1990). In this manner, then, the output of the surface phonetic form is inevitable. Even in situations that would favor, at least logically, the suppression of the generation of the surface form, it is nevertheless generated—at the cost of causing phonological confusion.

Although the findings of this study provide us with more evidence for the phonological nature of STM, especially for the inclusion of surface phonetic representation in STM, there are still important details about the process of STM that we do not yet understand well. For example, the modular account predicts the generation of both the underlying representation and the surface representation of linguistic material in STM. If this is true, would properties of the underlying representation also affect STM? For instance, would the span of STM be a function of the similarity in the underlying forms when the similarity in surface forms is held constant? Or, would the number of phonemes affect STM span when the surface duration of words is held constant? Finding answers to these questions would certainly further our knowledge of the phonological nature of short-term memory.

Finally, the finding that certain phonological rules are applied during the encoding of visually presented material by readers of a language with a nonalphabetic writing system also provides further evidence for the similarity of reading processes between languages with different writing systems. It is easy to believe that readers of languages with alphabetic writing systems access the pronunciation of words analytically, because the writing system more or less directly reflects the phonological structures of words. For a language with a nonalphabetic writing system, on the other hand, it is easy to believe that a "whole word" approach should be used; that is, the pronunciation is only activated through lexical access. However, there is evidence that even for languages with an alphabetic writing system, pronunciations of familiar words are activated through lexical access rather than through applying pronunciation rules to the spelling of the words (Baddeley & Lewis, 1981; Besner, Davies, & Daniels, 1981). To pronounce unfamiliar words, even Chinese words, on the other hand, readers have to resort to the phonetic component (if there is one) of the character to get the pronunciation (Seidenberg, 1985; Shu & Zhang, 1987), although frequently the guessed pronunciation is not the correct one. The finding of this study, especially that of Experiment 2, further suggests that even for a language like Chinese, the phonological structures at the word level could be analytical, although, perhaps, not cognitively analytical. The pronunciation of a sequence of monosyllabic morphemes is not necessarily a simple combination of the pronunciations of the component morphemes; rules have to be applied to derive the surface phonetic structure. Thus, rather than always being accessed directly, the pronunciation of the words may be derived through the application of pronunciation rules, especially for unfamiliar words. This is probably similar to the stress assignment in languages such as English. In this sense, there should be no basic difference in the process of lexical access for languages with a nonalphabetic writing system, such as Chinese, and for languages with an alphabetic writing system, such as English.

REFERENCES

Averbach, E., & Coriell, A. (1961). Short-term memory in vision. *Bell System Technical Journal*, **40**, 309-328.

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic and formal similarity. *Quarterly Journal of Experimental Psychology*, **18**, 362-365.

Baddeley, A. D. (1968). How does acoustic similarity influence short-term memory? *Quarterly Journal of Experimental Psychology*, **20**, 249-264.

Baddeley, A. D., & Hitch, G. B. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 199-241). New York: Academic Press.

BADDELEY, A. D., & LEWIS, V. J. (1981). Inner active processes in reading: The inner voice, the inner ear and the inner eye. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading* (pp. 107-129). Hillsdale, NJ: Erlbaum.

BADDELEY, A. D., THOMSON, N., & BUCHANAN, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **14**, 575-589.

BESNER, D., DAVIES, J., & DANIELS, S. (1981). Phonological processes in reading: The effects of concurrent articulation. *Quarterly Journal of Experimental Psychology*, **33**, 415-438.

BRADY, S., MANN, V., & SCHMIDT, R. (1987). Errors in short-term memory for good and poor readers. *Memory & Cognition*, **15**, 444-453.

CONRAD, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology*, **55**, 75-84.

CONRAD, R. (1972). Speech and reading. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye: The relationships between speech and reading* (pp. 205-240). Cambridge, MA: MIT Press.

ERICKSON, D., MATTINGLY, I. G., & TURVEY, M. (1977). Phonetic activity in reading: An experiment with Kanji. *Language & Speech*, **20**, 384-403.

FODOR, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

FRICK, R. W. (1985). Testing visual short-term memory: Simultaneous versus sequential presentations. *Memory & Cognition*, **13**, 346-356.

HUE, C.-W., & ERICKSON, J. R. (1988). Short-term memory for Chinese characters and radicals. *Memory & Cognition*, **16**, 196-205.

LIBERMAN, A. M., & MATTINGLY, I. G. (1985). The motor theory of speech perception revised. *Cognition*, **21**, 1-36.

LOCKE, J. L., & KUTZ, K. J. (1975). Memory for speech and speech for memory. *Journal of Speech & Hearing Research*, **18**, 176-191.

MATTINGLY, I. G., & LIBERMAN, A. M. (1990). Speech and other auditory modules. In G. M. Edelman, W. E. Gall, & W. M. Cowan (Eds.), *Signal and sense: Local and global order in perceptual maps* (pp. 501-520). New York: Wiley.

REN, N., & MATTINGLY, I. G. (in press). Short-term serial recall performance by good and poor readers of Chinese. *Haskins Laboratories Status Report on Speech Research*.

SEIDENBERG, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition*, **19**, 1-30.

SHANKWEILER, D., LIBERMAN, I. Y., MARK, L. S., FOWLER, C. A., & FISHER, F. W. (1979). The speech code and learning to read. *Journal of Experimental Psychology: Human Learning & Memory*, **5**, 531-545.

SHU, H., & ZHANG, H.-C. (1987). The process of pronouncing Chinese characters by proficient mature readers. *Acta Psychologica Sinica*, **19**, 282-290.

TZENG, O. J. L., HUANG, D. L., & WANG, W. S.-Y. (1977). Speech recoding in reading Chinese characters. *Journal of Experimental Psychology: Human Learning & Memory*, **3**, 621-630.

WANG, W. S.-Y., & LI, K.-P. (1967). Tone 3 in Pekinese. *Journal of Speech & Hearing Research*, **10**, 629-636.

WAUGH, H. C., & NORMAN, D. H. (1965). Primary memory. *Psychological Review*, **72**, 89-104.

WICKELGREN, W. A. (1965a). Distinctive features and errors in short-term memory for English consonants. *Journal of the Acoustical Society of America*, **39**, 388-398.

WICKELGREN, W. A. (1965b). Distinctive features and errors in short-term memory for English vowels. *Journal of the Acoustical Society of America*, **38**, 583-588.

XIA, Z.-N. (ED.). (1979). *Cihai* [A grand dictionary of Chinese words]. Shanghai: Shanghai Dictionary Press.

YIK, W. F. (1978). The effect of visual and acoustic similarity on short-term memory for Chinese words. *Quarterly Journal of Experimental Psychology*, **30**, 487-494.

ZHANG, G. & SIMON, H. A. (1985). STM capacity for Chinese words and idioms: Chunking and acoustical loop hypotheses. *Memory & Cognition*, **13**, 193-201.

NOTES

1. The significant effect taken to be that of visual similarity in Yik (1978) might have been confounded by some other factors. Of the five pairs of monosyllabic words considered to contain high visual similarity and low acoustic similarity, four pairs have either the same consonant (e.g., *shuai* vs. *shi*), or the same glide (e.g., *xuan* vs. *lu*), and one pair is semantically similar (*fen*: powder vs. *li*: small particle). In the inventory with low visual and high acoustic similarity, which turned out to have produced fewer errors than the inventory with high visual and high acoustic similarity, two pairs are not acoustically identical (e.g., *zhi* vs. *zi*).

2. Incidentally, their homophone set contains a character that has alternative tones, that is, *gōng* and *gòng*.

3. At the word level, there is a fifth tone, which is usually called the neutral tone. This tone occurs only after a full lexical tone within a word.

4. Mandarin, or Putonghua (the Common Chinese), as it is called in mainland China, or Guoyu (the National Language), as it is called in Taiwan, is an official language based on the Beijing dialect and other northern dialects of China. It is used mainly on radio, TV, and on other public occasions.

(Continued)

APPENDIX A
Sample Lists of Monosyllabic Morphemes (a) with Same Tones and (b) with Different Tones

(a)

1)	班 bān	肝 gān	安 ān	贪 tān	山 shān	翻 fān	三 sān
2)	残 cán	谈 tán	篮 lán	然 rán	凡 fán	含 hán	男 nán
3)	产 chǎn	胆 dǎn	砍 kǎn	伞 sǎn	反 fǎn	敢 gǎn	展 zhǎn
4)	烂 làn	饭 fàn	慢 màn	叛 pàn	站 zhàn	岸 àn	汉 hàn
5)	粗 cū	姑 gū	屋 wū	租 zū	书 shū	出 chū	突 tū
6)	读 dú	胡 hú	奴 nú	图 tú	如 rú	除 chú	吴 wú
7)	府 fǔ	苦 kǔ	母 mǔ	主 zhǔ	普 pǔ	鲁 lǔ	补 bǔ
8)	布 bù	树 shù	路 lù	度 dù	促 cù	护 hù	速 sù

(b)

1)	肝 gān	残 cán	展 zhǎn	山 shān	慢 màn	然 rán	反 fǎn
2)	产 chǎn	凡 fán	叛 pàn	班 bān	含 hán	岸 àn	贪 tān
3)	汉 hàn	谈 tán	砍 kǎn	三 sān	敢 gǎn	翻 fān	烂 làn
4)	男 nán	站 zhàn	胆 dǎn	安 ān	饭 fàn	篮 lán	伞 sǎn
5)	粗 cū	度 dù	吴 wú	补 bǔ	如 rú	书 shū	普 pǔ
6)	主 zhǔ	树 shù	护 hù	除 chú	突 tū	租 zū	读 dú
7)	路 lù	胡 hú	出 chū	速 sù	屋 wū	府 fǔ	苦 kǔ
8)	奴 nú	图 tú	促 cù	姑 gū	鲁 lǔ	母 mǔ	布 bù

APPENDIX B
Test Lists for Experiment 2

Sandhi List

1.	常肯 chángkěn	肠者 chángzhě	厂把 chǎngbǎ
2.	迟永 chíyǒng	齿法 chǐfǎ	池给 chígěi
3.	楚品 chǔpǐn	除甲 chújiǎ	厨本 chúběn
4.	宇也 yǔyě	语广 yǔguǎng	余久 yújiǔ
5.	府伙 fǔhuǒ	浮北 fúběi	抚瓦 fǔwǎ
6.	级果 jíguǒ	几可 jǐkě	挤凯 jǐkǎi
7.	节免 jiémiǎn	杰挺 jiétǐng	姐否 jiěfǒu
8.	兰纽 lánniǔ	览整 lǎnzhěng	蓝引 lányǐn

9.	礼丙 lǐbǐng	厘走 lízǒu	梨写 líxiě
10.	止小 zhǐxiǎo	纸吕 zhǐlǚ	执景 zhíjǐng
11.	每拐 měiguǎi	眉考 méikǎo	美总 měizǒng
12.	其忍 qírěn	启朵 qǐduǒ	企铁 qǐtiě
13.	逐举 zhújǔ	竹远 zhúyuǎn	煮冷 zhǔlěng
14.	石怎 shízěn	史顶 shǐdǐng	实很 shíhěn
15.	统角 tǒngjiǎo	童两 tóngliǎng	铜左 tóngzuǒ
16.	碗准 wǎnzhǔn	晚轨 wǎnguǐ	完表 wánbiǎo
17.	网吵 wǎngchǎo	亡妥 wángtuǒ	往短 wǎngduǎn
18.	维选 wéixuǎn	伪损 wěisǔn	尾且 wěiqiě
19.	无软 wúruǎn	吴捆 wúkǔn	午纽 wǔniǔ
20.	习缓 xíhuǎn	洗逞 xǐchěng	席检 xíjiǎn
21.	显耳 xiǎněr	闲孔 xiánkǒng	咸仅 xiánjǐn
22.	演所 yǎnsuǒ	掩响 yǎnxiǎng	严口 yánkǒu
23.	以扫 yǐsǎo	移暖 yínuǎn	椅柳 yǐliǔ
24.	尤审 yóushěn	友粉 yǒufěn	有董 yǒudǒng

No-Sandhi List

1.	常并 chángbìng	肠日 chángrì	厂寸 chǎngcùn
2.	迟后 chíhòu	齿过 chǐguò	池页 chíyè
3.	楚问 chǔwèn	除令 chúlìng	厨运 chúyùn
4.	宇或 yǔhuò	语克 yǔkè	余动 yúdòng
5.	府向 fǔxiàng	浮备 fúbèi	抚困 fǔkùn
6.	级认 jírèn	几正 jǐzhèng	挤定 jǐdìng
7.	节最 jiézuì	杰命 jiémìng	姐众 jiězhòng
8.	兰奋 lánfèn	览彻 lǎnchè	蓝次 láncì

9.	礼况 lǐkuàng	厘序 líxù	梨这 lízhè
10.	止共 zhǐgòng	纸列 zhǐliè	执外 zhíwài
11.	每壮 měizhuàng	眉册 méicè	美劝 měiquàn
12.	其会 qíhuì	启化 qǐhuà	企罢 qǐbà
13.	逐竟 zhújìng	竹幸 zhúxìng	煮样 zhǔyàng
14.	石对 shíduì	史面 shǐmiàn	实律 shílǜ
15.	统近 tǒngjìn	童够 tónggòu	铜念 tóngniàn
16.	碗甚 wǎnshèn	晚个 wǎngè	完落 wánluò
17.	网要 wǎngyào	亡料 wángliào	往碰 wǎngpèng
18.	维秀 wéixiù	伪顿 wěidùn	尾用 wěiyòng
19.	无下 wúxià	吴透 wútòu	午欠 wǔqiàn
20.	习大 xídà	洗宋 xǐsòng	席去 xíqù
21.	显断 xiǎnduàn	闲旧 xiánjiù	咸社 xiánshè
22.	演庙 yǎnmiào	掩庆 yǎnqìng	严创 yánchuàng
23.	以贵 yǐguì	移配 yípèi	椅灭 yǐmiè
24.	尤论 yóulùn	友扣 yǒukòu	有赤 yǒuchì

(Manuscript received June 11, 1990;
revision accepted for publication October 29, 1990.)