Use of Partial Information in Learning to Read Chinese Characters

Richard C. Anderson, Wenling Li, and Yu-Min Ku
University of Illinois at Urbana–Champaign

Hua Shu and Ningning Wu
Beijing Normal University

This study investigated whether children can use partial information to learn the pronunciations of Chinese characters. Participants were 49 2nd graders and 56 4th graders whose home language was Mandarin and 75 2nd graders and 93 4th graders whose home language was Cantonese. Children had 2 trials to learn the Mandarin pronunciations of 28 unfamiliar compound characters of 4 types. Children learned to pronounce more regular characters, which contain full information about pronunciation, and more tone-different and onset-different characters, which contain partial information about pronunciation, than characters with unknown phonetic components, which contain no information about pronunciation. Mandarin-speaking children learned more pronunciations than Cantonese-speaking children.

A notable feature of modern standard Chinese, or Mandarin, is that it contains only about 400 syllables. When further differentiated according to one of four tones or voice inflections—high, rising, low then rising, or falling—there are about 1,200 distinct “tone syllables” (Taylor, 2001). A Mandarin syllable consists of an optional initial consonant (traditionally called the initial; we will use the term onset) followed by a final element (called the final or rime), consisting of an obligatory nuclear vowel and one or more optional vowels and an optional consonant. Mandarin syllables can be described in terms of three phonological elements: onset, rime, and tone. All of the valid syllables of Mandarin can be defined in terms of the four tones and just 21 onsets and 36 rimes.

Every writing system connects with spoken language in some way. But unfortunately for Chinese children trying to learn to read, the connections in Chinese are indirect, complicated, and notoriously unreliable. The pronunciation of about 28% of the 2,570 characters in school Chinese is entirely arbitrary (Shu, Chen, Anderson, Wu, & Xuan, 2002). Characters with arbitrary pronunciations include all of the many simple characters children are supposed to learn in elementary school as well as a small number of semantic compound characters that contain information about meaning but not pronunciation. The remaining 72% of characters in school Chinese are semantic–phonetic compounds (sometimes called ideophonic compounds or, simply, phonetic compounds). Semantic–phonetic compound characters contain a semantic component that gives a clue to meaning and a phonetic component that may give a clue to pronunciation. The problem is that the clues to pronunciation in phonetic components are not very trustworthy. Only 23% of the compound characters prescribed to be taught during the six grades of elementary school by the Chinese Ministry of Education are perfectly regular in the sense that the whole compound has exactly the pronunciation expected on the basis of its phonetic component (Shu et al., 2002). However, a large additional group of compound characters in school Chinese contain a phonetic component that provides partial information about the pronunciation of the whole compound. This article reports an empirical investigation of whether Chinese children can use partial information of two different kinds to learn and remember the pronunciations of unfamiliar compound characters.

A notable feature of modern standard Chinese, or Mandarin, is that it contains only about 400 syllables. When further differentiated according to one of four tones or voice inflections—high, rising, low then rising, or falling—there are about 1,200 distinct “tone syllables” (Taylor, 2001). A Mandarin syllable consists of an optional initial consonant (traditionally called the initial; we will use the term onset) followed by a final element (called the final or rime), consisting of an obligatory nuclear vowel and one or more optional vowels and an optional consonant. Mandarin syllables can be described in terms of three phonological elements: onset, rime, and tone. All of the valid syllables of Mandarin can be defined in terms of the four tones and just 21 onsets and 36 rimes.

A major class of characters that provides partial information about pronunciation can be called tone-different because each character in the class is pronounced with the same onset and rime (and thus the same syllable) but a different tone from its phonetic component. For example, the character 序 is pronounced /qıŋ/, whereas its phonetic 序 is pronounced /qıŋ/. About 16% of the semantic–phonetic compound characters in school Chinese are tone different (Shu et al., 2002).

Tone-different characters are usually classified as regular (e.g., Tzeng, Zhang, Hung, & Lee, 1995). Presumably this classification reflects the intuition of adult readers of Chinese that they can easily make use of the information about onset and rime in the phonetic of tone-different characters. However, we are aware of just one previous study that examined whether children can use the partial information about pronunciation potentially available in tone-different characters. Ho and Bryant (1997) presented familiar semantic–phonetic compound characters to Hong Kong first and second graders. They found that completely regular compound characters (same onset, rime, and tone as the phonetic) were read significantly better than tone-different characters, which in turn were read significantly better than irregular characters (different onset, rime, and tone). Pronunciation of tone-different characters...

Richard C. Anderson, Wenling Li, and Yu-Min Ku, Center for the Study of Reading, University of Illinois at Urbana–Champaign; Hua Shu and Ningning Wu, Department of Psychology, Beijing Normal University, Beijing, China.

We are pleased to acknowledge that the research described herein was supported by the Spencer Foundation. We wish to thank Xi Chen for her contributions to the research and Aydin Durgunoglu for her comments on a draft of this article.

Correspondence concerning this article should be addressed to Richard C. Anderson, Center for the Study of Reading, University of Illinois at Urbana–Champaign, 174 Children’s Research Center, 51 Gerty Drive, Champaign, Illinois 61820. E-mail: csrrca@uiuc.edu

1 Pronunciations are enclosed in back slashes. The diacritical marks represent tones, for example, /má/, /má/, /má/, and /má/ for the first through fourth tones, respectively.
was conditioned by frequency. Among high-frequency characters, tone-different characters were read as well as regular characters, whereas among low-frequency characters, tone-different characters were read no better than irregular characters. From this Ho and Bryant concluded, “Only from repeated exposure to [tone-different] characters were . . . [children] able to associate the sound of the phonetic component with that of the characters as a whole” (p. 283).

Another large class of compound characters that provides partial information about pronunciation can be called *onset different*. These are characters pronounced with a different onset but the same rime as the phonetic component. For instance, 北 is pronounced *běi*, whereas its phonetic 北 is pronounced *lìng*. Shu et al. (2002) calculated that 20% of the semantic–phonetic compound characters in school Chinese are onset different. There is no previous research with either adults or children that addresses whether readers are able to use the information about rime contained in onset-different characters.

Tone-different and onset-different characters together account for 36% of the semantic–phonetic compounds in school Chinese (Shu et al., 2002). Thus, the ability to use partial information about pronunciation would be extremely useful to young readers. This article reports an empirical investigation of whether they are able to do so.

The theoretical notion that guides this research is that learning to read entails understanding the principles that govern the relationships between speech and writing. In English, the overarching principle is that letters represent segmental phonemes in spoken English (Liberman, Shankweiler, & Liberman, 1989; Treiman, 2000). The alphabetic principle motivates and rationalizes the assimilation of specific letter–sound associations for the young reader. However, to become skilled readers of English, children have to acquire a more sophisticated understanding of orthography–phonology relationships as they learn to cope with variability in the pronunciation of letters and letter clusters (Goswami, 2000).

In Chinese, the overarching graphophonological insight might be called the *phonetic principle* (in Shu, Anderson, & Wu, 2000, we used the term *phonetic awareness*). The phonetic principle is simply that the phonetic components of compound characters represent syllables in spoken Chinese. Young readers—and older, poor readers—are not necessarily aware of, or functionally able to use, this principle (Ho & Bryant, 1997; Shu et al., 2000); they may encode characters as a whole, in which case the pronunciation of compound characters is arbitrary and pronunciations must be memorized one by one. Shu et al. (2000) pointed out that, at first, a child’s awareness of the phonetic principle may be unsophisticated, perhaps, if the child were able express it: “The part on the right tells the pronunciation” (p. 61). They supposed that a more sophisticated view emerges during elementary school, a view which “presumably incorporates insights into the partial information available in semi-regular characters” (p. 62).

One type of evidence that mastery of the alphabetic principle is critical for learning to read English and other alphabetic languages is the high correlation of pseudoword reading with word recognition and reading comprehension (e.g., Siegel, 1993), where the pseudowords represent monosyllables without obscure or tricky letter–sound relationships. Comparable, although still preliminary, evidence has been found in support of the importance of the phonetic principle in learning to read Chinese. Ho and Bryant (1997) asked Hong Kong first and second graders to pronounce compound characters, two-character words, and pseudocharacters. The pseudocharacters were novel combinations of familiar semantic radicals and familiar phonetic components in their legal positions. Pseudocharacter pronunciation was found to correlate highly with the pronunciation of both compound characters and two-character words. This implies that children who are aware of and attempting to use the information in the phonetic component are making better progress in learning to read. Similarly, C. K. Chan and Siegel (2001) had Hong Kong children in the first through the sixth grade read aloud a list of characters, about 80% of which were semantic–phonetic compounds. Young normal readers had significantly higher scores on a pseudocharacter pronunciation task than young poor readers. These findings again imply that better readers are trying to use the information in the phonetic component.

That pseudocharacter pronunciation correlates highly with the pronunciation of real Chinese characters and words creates a paradox. Pronouncing a compound pseudocharacter involves no more than naming its phonetic component. Yet, simply naming the phonetic is not a strategy that can work well when pronouncing real characters because of the low regularity and consistency of real characters. To reiterate, only 23% of the compound characters in school Chinese are entirely regular. It is apparent, therefore, that a Chinese child cannot reliably predict the pronunciations of unfamiliar compound characters. Nonetheless, the partial information about pronunciation available in many less-than-fully regular characters may be helpful for assimilating the characters, once the pronunciation has been provided by the teacher, looked up in the dictionary, or figured out from context while reading.

Ehri (1991, p. 402) has illustrated the use of partial information for assimilating English words, using the example *sword*. Although the word is irregular, and an English-speaking child who did not know the word would not be able to predict its pronunciation or spelling, the letters *s*, *o*, *r*, and *d* have their usual sound values. Ehri’s research shows that the phonemically aware child can use this partial information to readily encode and remember *sword* and other English words containing partial information. In a comparable way, we hypothesized that Chinese children can use the partial information in tone-different and onset-different characters to encode and remember the characters. This hypothesis was evaluated in the present study.

The research reported in this article also evaluated the influence of home language on learning to pronounce characters. Distinct dialects of Chinese, so different that they are actually mutually unintelligible languages, are spoken in the various regions of China. Nevertheless, Mandarin is the language of instruction everywhere. The children who participated in the present study were from Beijing, where the home language of children is Mandarin, and from Guangzhou, where the home language of most children is Cantonese. Cantonese-speaking children learn to speak, read, and write Mandarin when they enter primary school. From the first grade on, only Mandarin is supposed to be used in the classroom. Children are admonished to speak Mandarin in the hallways and on the playground.

---

2 This total includes characters that vary in tone as well as onset.
Cantonese syllables, like those of Mandarin, consist of an onset, a rime, and a tone. However, pronunciations have diverged over the past 2,000 years. For example, whereas Mandarin has four tones, Cantonese is considered to have nine. Many words in the two dialects are cognates, that is, related in meaning and with related but distinct pronunciations, much the same as the words *reason* and *raison*, which are cognates in English and French, respectively. These differences and similarities should influence Cantonese-speaking children learning to read Mandarin in several ways. In this study, we investigated whether the differences between home languages affected children’s ability to use partial information to learn character pronunciations. It would not be surprising to find that bilingual children whose home language is Cantonese are slower than Mandarin-only children to learn Mandarin pronunciations. However, we did not undertake the study committed to a strong directional hypothesis. Previous research indicates that bilingual children may develop superior metalinguistic awareness that may help compensate for lack of fluency or specific interference (Bialystok, 1986, 2001; Durgunoglu & Verhoeven, 1998).

To recapitulate, we addressed three questions: (a) Can Chinese children use partial information to help learn and remember the pronunciations of unfamiliar tone-different and onset-different characters? (b) Which kind of partial information can children most readily use? (c) Does home language influence children’s ability to learn pronunciations of unfamiliar characters?

**Method**

**Participants**

The participants were Spring semester second graders and fourth graders from comparable working-class schools in Beijing and Guangzhou. There were 49 second graders and 56 fourth graders from Beijing and 75 second graders and 93 fourth graders from Guangzhou. Parents from Guangzhou completed a questionnaire about the language used in the home; we excluded students whose home language was not Cantonese. There were three between-participants factors, home language (Mandarin, Cantonese), grade (second, fourth), and teacher (low, middle, high) as well as two within-participants factors, type of character (four types) and trial (two trials). Children had two trials to learn the pronunciations of 28 unfamiliar characters. During each trial, an experimenter standing in the front of the classroom displayed and pronounced the characters one by one in a random order at about a 2-s rate. Then the children indicated the pronunciation of each character on an answer sheet on which the 28 characters appeared in a different random order. The method of presenting and testing the characters was the same during the second trial, except that the characters were in different random orders.

The children indicated pronunciations by writing pinyin. Pinyin is an alphabetic script children in the People’s Republic of China learn in the first 10 weeks of first grade. The purpose of pinyin, literally “spell sound,” is to help children learn the pronunciations of characters. All of the characters in the first-grade textbooks are annotated with pinyin, and pinyin is provided for the new characters in the second- and third-grade textbooks. Having children write pinyin is an inexpensive option to having them actually pronounce characters. An additional advantage of using pinyin in the present study is that it side steps ambiguities in oral pronunciation, which might be acute among dialect speakers. Children are highly familiar with pinyin and easily able to use it. Meng, Shu, and Zhou (2000) found a correlation of .93 between oral pronunciations and pronunciations written in pinyin.

Four types of semantic–phonetic compound characters were used: (a) regular characters, in which the phonetic component provides full information (onset, rime, and tone) about the pronunciation of the character; (b) tone-different characters, in which the phonetic has the same onset and rime but not the same tone; (c) onset-different characters, in which the phonetic has the same rime and tone but not the same onset; and (d) characters with phonetics that children do not know, which therefore do not convey any information about pronunciation. There were seven unfamiliar characters of each type, unfamiliar in the sense that none of them were listed to be explicitly taught in the textbooks for the first, second, third, or four grades. The characters of each type were equated in frequency and in number of strokes. All the characters had a left–right structure.

Teachers rated children’s reading level. These ratings served to classify the children as low, middle, or high for the data analysis. Children also completed grade-appropriate reading comprehension tests. However, with one exception, these tests contained only a few items, had low reliability, and will not be discussed further. The exception was the second-grade sentence comprehension test, which consisted of 25 sentences (based on a test by Nielsen, 1993). Each sentence was accompanied by four pictures. The children’s task was to select the picture that best represented the meaning of the sentence. This test had an estimated lower bound reliability of .77 (Cronbach’s alpha).

**Results**

The major finding of the study was that type of character strongly influenced the learning of pronunciations, Wilks’s $\lambda = .123$, $F(3, 235) = 556.88, p < .01$, $\eta^2_p = .877$. The high value of partial eta squared ($\eta^2_p$) confirms a very large effect size; nearly 88% of between-characters sample variance is explained by type of character. Planned comparisons indicate that, as expected, performance was better on regular, $t(207) = 36.09, p < .01$, and tone-different, $t(207) = 19.76, p < .01$, characters than characters with unknown phonetics. As can be seen in Table 1, the advantage of onset-different characters over phonetic-unknown characters was much smaller, although as forecast this difference, too, was significant, $t(207) = 3.56, p < .01$.

A second major finding was that monolingual children whose home language is Mandarin were better able to learn the Mandarin pronunciations of unfamiliar characters than bilingual children whose home language is Cantonese, $F(1, 237) = 22.70, p < .01$, $\eta^2_p = .087$. A significant interaction was found between character type and home language, Wilks’s $\lambda = .938$, $F(3, 235) = 5.17, p < .01$, $\eta^2_p = .062$. As Figure 1 shows, there were differences between monolingual and bilingual students on regular, tone-different, and onset-different characters but not on characters with unknown phonetics. This result indicates that monolingual children are better able than bilingual children to use the information provided by the phonetic component.

Children averaged 23% correct pronunciations on the first trial and 36% on the second trial, which is a significant improvement, Wilks’s $\lambda = .316$, $F(1, 237) = 513.16, p < .01$, $\eta^2_p = .684$. Grade had a significant effect, $F(1, 237) = 93.78, p < .01$, $\eta^2_p = .284$. Fourth graders could pronounce 35% of the characters, whereas second graders could pronounce 23%. Neither the Character Type $\times$ Grade, Grade $\times$ Home Language, nor the Character Type $\times$ Grade $\times$ Home Language interactions were significant.
Reading level, as rated by teachers, significantly influenced performance, $F(2, 237) = 62.14, p < .01, \eta^2_p = .344$. Moreover, there was a Character Type $\times$ Reading Level interaction, Wilks’s $\Lambda = .935, F(6, 472) = 2.69, p < .02, \eta^2_p = .033$. Table 2 displays the percentage of correct pronunciations by character type and reading level. Reading level is less strongly associated with performance on regular characters than the characters of the other three types. Also significant was the Character Type $\times$ Reading Level $\times$ Grade interaction, Wilks’s $\Lambda = .918, F(6, 472) = 3.43, p < .01, \eta^2_p = .041$. This interaction qualifies the two-way interaction just described, indicating that the effect is mainly due to fourth graders: Low and average fourth graders perform almost as well as high fourth graders on regular characters, whereas reading level is more strongly related to performance on the other three types of characters in the fourth grade and every type of character in the second grade.

An analysis of children’s errors in reproducing the pronunciation of the newly learned characters is summarized in Table 3. Errors consist of mistakes or omissions on one or more of the three phonological features for describing Chinese pronunciations—onset, rime, and tone. For example, \texttt{RÈ /bā/} is a regular character.

If a student wrote /pā/, then an onset error was tallied in the regular character column; if the student wrote /pā/, then a tone error as well as an onset error was tallied. Both second and fourth graders made more errors when less information was available in the phonetic. That is, children made progressively more errors on regular, tone-different, onset-different, and phonetic-unknown characters. Not surprisingly, when the phonetic component provided partial information, children made more errors on the element of the compound character that was not consistent with its phonetic component. That is, they made more tone errors on tone-different characters and more onset errors on onset-different characters. When the phonetic component provided full information or no information about the pronunciation of the whole character, the errors children made were more or less evenly distributed across the three phonological elements.

Cantonese-speaking second graders averaged 83% correct on the sentence comprehension test, whereas Mandarin-speaking second graders averaged 78% correct, which is not a significant difference. For second graders, the correlations of sentence comprehension scores and teacher’s rating of reading level with total characters correctly pronounced were .11 and .48 ($p < .01$) for Cantonese speakers and .17 and .56 ($p < .01$) for Mandarin speakers, respectively. For fourth graders, the correlations of teacher’s rating of reading level with total characters correctly pronounced were .52 ($p < .01$) for Cantonese speakers and .55 ($p < .01$) for Mandarin speakers.

Table 2

<table>
<thead>
<tr>
<th>Type of character</th>
<th>Low</th>
<th>Middle</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$SD$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regular</td>
<td>48</td>
<td>27</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>58</td>
<td>25</td>
<td>69</td>
</tr>
<tr>
<td>Tone different</td>
<td>23</td>
<td>18</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>21</td>
<td>48</td>
</tr>
<tr>
<td>Onset different</td>
<td>8</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>14</td>
<td>21</td>
</tr>
<tr>
<td>Phonetic unknown</td>
<td>5</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>All types</td>
<td>21</td>
<td>19</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>19</td>
<td>38</td>
</tr>
</tbody>
</table>

Figure 1. Percentage correct pronunciations as a function of character type and home language.
Table 3
Proportions of Errors as a Function of Character Type

<table>
<thead>
<tr>
<th>Type of error</th>
<th>Regular</th>
<th>Tone different</th>
<th>Onset different</th>
<th>Unknown phonetic</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Onset error</td>
<td>.39</td>
<td>.44</td>
<td>.88</td>
<td>.94</td>
</tr>
<tr>
<td>Rime error</td>
<td>.38</td>
<td>.51</td>
<td>.58</td>
<td>.96</td>
</tr>
<tr>
<td>Tone error</td>
<td>.45</td>
<td>.64</td>
<td>.65</td>
<td>.91</td>
</tr>
<tr>
<td>4th grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Onset error</td>
<td>.20</td>
<td>.25</td>
<td>.73</td>
<td>.75</td>
</tr>
<tr>
<td>Rime error</td>
<td>.22</td>
<td>.30</td>
<td>.35</td>
<td>.75</td>
</tr>
<tr>
<td>Tone error</td>
<td>.26</td>
<td>.49</td>
<td>.43</td>
<td>.72</td>
</tr>
</tbody>
</table>

Discussion

This study convincingly demonstrates that children as young as second grade can make use of information in the phonetic component to learn the pronunciations of novel compound characters. Especially newsworthy is the fact that they can use partial information about pronunciation. Of the two types of characters examined in the present study that contain partial information, the information in tone-different characters proved to be readily accessible to children, whereas the information in onset-different characters was marginally accessible to them.

The present study confirms and extends the finding of Ho and Bryant (1997) that children can use the partial information about pronunciation in tone-different characters. However, both studies also demonstrate the limits of partial information: performance on tone-different characters is substantially worse than performance on completely regular characters, which of course provide full information about pronunciation. During the second trial in the present study, neither second graders nor fourth graders had come up to the level of performance attained during the first trial on completely regular characters. As Ho and Bryant said, repeated exposure is necessary for children to learn the pronunciation of tone-different characters.

The finding that children can use the information in onset-different characters is without precedent in the previous literature. It is doubly true, however, that repeated exposure is necessary to learn onset-different characters. During the second trial, children were still far worse on onset-different characters than during the first trial on either regular or tone-different characters.

Some level of metalinguistic awareness is implicated as a prerequisite for a child to make use of partial information about pronunciation (Nagy & Anderson, 1998; Shu & Anderson, 1999; Treiman & Zukowski, 1991). Specifically, tone awareness would seem to be necessary to make use of the information in tone-different characters, and onset–rime awareness would seem to be necessary for children to use the information in onset-different characters. By tone awareness and onset–rime awareness, we mean the ability to identify, reflect on, and manipulate these facets of language. Tone awareness, for example, can be assessed in an odd-man-out task. Among /hán/, /jiàng/, and /dōu/, for instance, the odd one is /hán/ because it has a different tone than the other two.

Previous research shows that facets of phonological awareness are important for learning to read Chinese (Hu & Catts, 1998; Huang & Hanley, 1995; Taylor, 2001). McBride-Chang and Ho (2000) found that syllable deletion was related to character recognition. So and Siegel (1997) reported a significant difference between poor and average readers in tone discrimination. Ho and Bryant (1997) found significant correlations of rime detection with character and pseudolibrary pronunciation. Li, Anderson, Nagy, and Zhang (2001) found significant relationships between syllable reversal, onset–rime discrimination, and especially tone discrimination with measures of reading. However, none of these studies evaluated the specific hypothesis that tone awareness and onset–rime awareness would facilitate learning the pronunciation of tone-different and onset-different characters, nor was the hypothesis evaluated in this study.

Another newsworthy finding from the present research is that children from Guangzhou, who are learning Mandarin as a second language, are less able to use phonetic information to learn Mandarin pronunciations than children from Beijing, who are monolingual speakers of Mandarin. Presumably children from Guangzhou performed well because of interference from Cantonese or lack of fluent access to Mandarin pronunciations, although bilingualism is complicated (see, e.g., Durgunoglu, Nagy, & Hancin-Bhatt, 1999) and we are unable to say exactly what the mechanism is. Despite relatively poor performance on the pronunciation task, Cantonese-speaking children performed as well as or better than Mandarin-speaking children on the measures of reading comprehension. This suggests that in Chinese there is not a tight coupling between ability to pronounce characters and ability to comprehend.

Fourth graders learned more pronunciations than second graders. Perhaps this implies that fourth graders have learned how to learn the pronunciations of unfamiliar characters. Or, maybe fourth graders knew the pronunciations of some of the supposedly unfamiliar characters, inasmuch as unfamiliar characters were operationally defined as characters not yet taught in school. Fourth graders may have figured out the pronunciations of some characters not yet taught in school when they encountered the characters in independent readings, or they may have looked up some of these characters in a dictionary.

The teacher’s rating was the only measure of reading level that had consistently high correlations with character pronunciation. This suggests that in China as well as in the United States (Hoge & Coladarci, 1989), teachers are a source of valid and reliable information about reading.

With the exception of the teacher rating, the correlations between the measures of reading and character pronunciation were mostly low. Some of the measures of reading contained small numbers of items and were unreliable. However, this was not true of the sentence comprehension measure used in the second grade, which had 25 items and reliability conservatively estimated at .77. Yet, this measure correlated only .11 and .17 with total characters correctly pronounced among Cantonese-speaking and Mandarin-speaking children, respectively. In learning to read Chinese, apparently there is a disjunction between comprehension and pronunciation. Children probably are able to use the information about meaning in the semantic radicals of compound characters (L. Chan & Nunes, 1998; Shu & Anderson, 1997) and their knowledge of word structure and syntax (Ku & Anderson, 2000; Li et al., 2001) to comprehend sentences well enough to match the sentences with pictures, without necessarily being able to accurately pronounce every character.
In conclusion, when we take a close look, awareness of the phonetic principle is not as rudimentary as it may appear to be among beginning readers. At the beginning stage, children’s understanding of the principle might be something like, “The part on the right tells the pronunciation” (Shu et al., 2000, p. 61). At this stage, children recognize the phonetic component of a semantic-phonetic character and use it directly to read the character. As they become more experienced in reading, children who are making good progress learn how to cope with more complicated relationships between the phonetic component and the pronunciation of a character. Among the things they learn is how to make use of partial information to learn and remember the pronunciations of novel characters.

Chinese teachers typically do not explain the logic of characters to their students (Wu, Li, & Anderson, 1999). They evidently feel it would be useless to explain the role of the phonetic component in compound characters, because such a large proportion of compound characters are irregular. The present study suggests that typical practice may be too conservative. Although we won’t know for sure until the research is done, it seems probable that if teachers were to point out the partial information about pronunciation available in semiregular characters, they might enable average- and low-performing students to use information that, to some extent, high-performing students have discovered how to use without special instruction.

References

ness in Chinese and English. Champaign, IL: Center for the Study of Reading.

Revision received April 30, 2002
Accepted May 2, 2002

57

PARTIAL INFORMATION IN LEARNING CHINESE CHARACTERS