1

Phonological Awareness, Orthography, and Learning to Read Chinese

Jun-Ren Lee and Chu-Ren Huang

1.1 Introduction

Phonological awareness is one of the most significant findings in research on reading in the last century (e.g., Mattingly 1972). The effect of phonological awareness has been attested by correlation studies, cross-sectional studies, longitudinal studies, and experimental intervention studies. These studies consistently show that phonological awareness is strongly correlated with reading acquisition, and in fact, it could be considered as the driving force of reading (Adams 1994). Interestingly, although phonological awareness is well attested in many languages and in most bilingual studies, the empirical data do not consistently support the relationship between phonological awareness and Chinese reading (Huang and Hanley 1995; Taylor 2002; Tzeng and Chen 曾世杰, 陳淑麗 2007). Given the inconsistent results, can they be attributed to the unique feature(s) of Chinese phonology, Chinese orthography, or other factors?

1.2 Theoretical Background

Research on phonological awareness and reading typically assume a simple processing model of reading with two basic components: word recognition (decoding) and language comprehension (Hoover and Gough 1990). The stages of reading development evolve from the stage of learning to read to the stage of reading to learn, with the transition occurring typically at the fourth grade. Learning to decode or sounding out a word is a critical factor in reading development (Chall 1983), while decoding a word is dependent on the development of phonological awareness and phonics (NICHD 2000). Previous studies on languages with alphabetic writing systems often assume that phonological awareness refers to the ability to represent and manipulate speech sound; and phonics the ability to segment the letter

string into grapheme, map it into the phoneme, and then synthesize the phonemes for the pronunciation of the word. Although these two aspects have been considered critical factors in early literacy (Snow et al. 1998), the theory clearly presupposes that writing systems represent phonological information and cannot be directly applied to languages with non-phonological writing systems.

Sproat's (2000) theory of writing systems introduced the concept of **orthographically relevant level** (ORL) of writing systems. He showed that writing systems of language typically take phonology as ORL, but not necessarily so. Several writing systems, including Egyptian hieroglyphs and Hanzi/kanji (Chinese characters) seem to have semantics as the ORL. As ORL indicates the linguistically relevant set of criteria for constructing a writing system, it predicts that this level of linguistic knowledge would be invoked during reading, as the process of accessing the linguistic and content knowledge represented by the text. This is the crucial theoretical underpinning of the theories for the cognitive process of reading but rarely spelled out explicitly. Underlining this theoretical underpinning also highlights the challenges of the research on phonological awareness in reading Chinese.

It is important to note that, other than reading, speakers' awareness of phonological knowledge can be verified with several different types of evidence. One of the most productive sources of evidence is the phonological adaptation of loanwords (e.g., LaCharité and Paradis 2005; Kenstowicz and Suchato 2006). Phonological adaptation requires the speakers to share the same awareness of their native language phonology in order to identify and adapt the non-complying part of the pronunciations of a loanword. Interestingly, literatures on phonological adaptation of loanwords in Mandarin Chinese (e.g., Hsieh and Kenstowicz 2008; Lin 2008), unlike reading, did not suggest any exceptional behavior.

Note that a premise that is shared with the process of phonological adaptation of loanwords with other phonologically relevant speech processes is that they are not dependent on orthography. Thus given that reading will necessarily involve orthography, a possible cause for the different phonological awareness results may be due to the differences in the writing systems. Thus, the exceptional results in Chinese reading are likely a reflection of orthographic awareness plus phonological awareness. In what follows, we will first introduce the basic theory of the phonology and orthography with special attention to the orthographically relevant level of Chinese. This will be followed by a section on other independent evidence of phonological awareness and orthographical awareness in Chinese, with some interesting data based on derivational phonology. The theoretical is followed by a comprehensive survey of psychological studies on reading and phonological awareness. Finally, the chapter concludes with a discussion and topics for future studies.

1.3 Chinese Phonology and Morphology: Theoretical Perspectives

A comprehensive introduction to Chinese phonology can be found in Duanmu (2007), which can be supplemented by the in-depth discussion of theoretical issues in the chapters in Part III of this volume (Duanmu 2022; Lin 2022; Neergaard and Huang 2022). For the purpose of discussion of phonological awareness, the most salient phonological features in Chinese include: (1) that the suprasegmental property of tone is phonologically relevant, (2) that the syllable inventory is relatively small with significant constraints such as the lack of consonant clusters and all close syllables end with nasal consonants, and (3) that tone sandhi applies to third tone and is context-sensitive.

To attest to phonological awareness, it is necessary to demonstrate that the speakers make use of phonological knowledge, either explicitly or implicitly. Thus we start our discussion with the kind of phonological knowledge that speakers have shown awareness of. We illustrate with the following examples that phonological awareness should include at least (1) syllabic identity, (2) tonal identity, (3) syllabicity, and (4) tonal value.

(1) (Dong and Wong 2020 (10c))
   a. 北大大夫 *běidà dài·fu* Peking_University-doctor ‘doctor of Peking University’
   b. 音樂樂園 *yīnyuè lèyuán* music-paradise ‘music paradise’
   c. 協會會計 *xiéhuì kuàijì* association-accountant ‘association’s accountant’
   d. 銀行行政 *yínháng xíngzhèng* bank-administration ‘bank administration’

The example above involves phonological conditions of haplology (Dong and Wong 2020). Haplology is one of the morpho-phonological features of Chinese, which was first described by Chao (1968) and can be generally defined as deletion of one of the two adjacent syllables under the conditions of identity. What is crucial is that semantic, orthographic, and phonological identity conditions must be met simultaneously (Huang and Wong 2018; Dong and Huang 董思聰, 黃居仁 2020; Dong and Wong 2020). For instance, haplology is blocked for 蒙牛牛奶 *méngniú niúnai* ‘Mengniu Milk’ even though both the orthographic and phonological identity conditions are met for the two consecutive *niú* syllables. This is because the semantic content of the two morphemes are different, *méngniú* referring to the company and not necessarily to any concept of cow. What (1) shows is that speakers are aware of syllabic boundaries. Haplology is blocked in all examples in (1) even though all of them involve two consecutive morphemes with identical written forms and even arguably identical semantics for (1a–b). This is because the condition of phonological identity *within the*

same syllable is not met. Thus the phonological awareness of syllabic structure for Mandarin speakers is attested even though such information is not reflected in the writing system.

Second, tone sandhi (Chao 1968; Yip 1980) has been the primary example used to illustrate the awareness of lexical tones at the phonological level for Mandarin speakers. That is, even though citation and sandhi tones are realized with different pitch contours, studies consistently show that speakers identify them as belonging to the identical lexical tone. This ability to cognitively 'ignore' perceivable differences has been considered to be the strongest evidence of phonological awareness.

Interestingly, haplology in the context of tone sandhi provides crucial evidence that speakers are also aware of the differences between sandhi vs. citation tones.

(2) 圖書館館長 *túshūguǎn guǎnzhǎng* 'head librarian'

It has been observed that haplology is often blocked when two identical third tone syllables occur consecutively. Mandarin speakers strongly dispprefer haplology for compounds such as (2). Given Chinese phonology, the obvious hypothesis is the non-identical tonal values due to tone sandhi blocking the phonological identity condition. This hypothesis is corroborated by recent comparative studies involving Hong Kong Chinese (Huang and Wong 2018; Wang and Dong 2020). The reported study showed that Cantonese speakers in Hong Kong tend to allow haplology in the context of tone sandhi, while Taiwan subjects are consistent in blocking. Since tone sandhi rules are non-obligatory in Cantonese, unlike Mandarin, it is likely that the differences of their awareness of sandhi form resulted in different preferences.

Lastly, syllable as segments is the most basic phonological concept that is widely assumed in studies of phonological awareness. However, it is also the hardest to prove independently, as most phonological processes involving a syllable also involve either the syllabic structure or the content of the syllable. One instance where only the concept of syllable is involved is Southern Min triple reduplication. In particular, the following gap involving Southern Min vivid reduplication reported in Huang (1992) is significant.

(3) a. 灵 *lêng* 'spiritually powerful'
b. * *lêng- lêng*
c. 灵灵灵 *lêng-lêng- lêng* 'spiritually very powerful'

In (3), it is shown that, contrary to expectation, triple reduplication of the morpheme *lêng* is attested while double reduplication is not. This distinction requires a speaker to be aware of the number of syllables. In addition, it is also discussed in both Cheng (1981) and Huang (1992) that the triple reduplication forms, such as (3c), in Southern Min can be reduced to only two syllables in production, with the condition that the combined contour

tone of the last two original syllables is superimposed on the second syllable of the abbreviated form. This provides further evidence showing the awareness of number of syllables by the speaker with the preservation of the third syllable suprasegmentally even though the segment itself is reduced.

The above discussion, complementing later chapters on phonology in this volume, showed that phonological awareness in Mandarin specifically, and in Sinitic languages in general, is not markedly different from other languages and that it does not presuppose any particular orthography. As reading involves both phonology and orthography, the theoretical issue has been at the center of past studies of Chinese reading, yet it is rarely explicitly expressed and is typically addressed without taking into consideration the linguistically relevant features of orthography.

1.3.1 Research Question: Do Specific Features of Chinese Orthography Influence Phonological Awareness in Reading?

The most salient feature of the Chinese writing system is that, unlike most writing systems in the world, phonology is not its orthographically relevant level. Sproat (2000) introduced the concept of ORL as the crucial link between writing and linguistic representation. He showed that phonology is the ORL for most of the writing systems of the world. It is important to note that the script employed is independent of the ORL. Hence, although languages can adapt Arabic, Cyrillic, Devanagari, Hangul, or Latin scripts (among others), they generally maintain the same ORL. One particularly illustrative example is Japanese, whose writing system allows the mixed codes of both hiragana and katakana, yet syllables are the same basic phonological units when either script is applied. Japanese also serves to underline that phonology as ORL can be applied at more than one level; the majority of alphabetic writing systems map to phoneme level, but the Japanese kana writing system is one of the best-known syllabic writing systems.

Note that *a priori*, Mandarin Chinese can as easily adopt an orthography with phonology as the ORL. In fact, both the Pinyin romanization and the bpmf national phonetic alphabet systems (bpmf NPA hereafter) are writing systems with phonology as ORL. Both of them encode phonemes as the basic writing units and reflect the awareness of the phonological knowledge discussed above. However, neither system was adopted conventionally in daily usage; nor could either have as profound an influence on the Chinese language as the Chinese character writing system that has been used continuously for nearly 3,000 years. What is the ORL for the character writing system and does that have any implication for reading and phonological awareness?

Sproat (2000) proposed, and Huang (2009) and Huang and Hsieh (2015) provided systematic arguments for semantics to be the ORL for the Chinese character writing system. The most illustrative examples are the two related monomorphemic words in (4).

(4) a. 枇杷 *pípá* 'loquat'
b. 琵琶 *pípá* 'pipa'

Etymologically the two words 'loquat' and 'pipa(instrument)' are related loanwords, although historical data does not show clearly which word was attested in Chinese first. They both belong to the small (less than or around 100) disyllabic morphemes. That is, the characters 枇,杷,琵,琶 and the syllables they represent are not morphemes and do not carry independent meaning. In general, these segments cannot stand alone, other than when they are used as abbreviated forms of the full disyllabic morphemes. If the Chinese writing system codes any phonological information, these forms, only without meaning characters, should be the primary candidates of being interchangeable, or at least confusing. That is, writing forms such as *枇琶 or *琵杷 should either be acceptable or at least well-attested errors. However, neither forms were attested historically nor in modern corpora. In fact, both Sproat (2000) and Huang and Hsieh (2015) were able to show that one significant rule in the Chinese orthography is that the two characters representing a disyllabic morpheme must always share the same radical, i.e., 木 mù (for plants) for loquat and the double 玉 (musical instrument) for pipa. More examples of disyllabic roots in Chinese include 葡萄 *pútáo* 'grape', 蝴蝶 *húdié* 'butterfly', and 徘徊 *páihuái* 'to linger, to walk back and forth'. Chou and Huang (2010) and Huang et al. (2013) clearly show the ontological system behind the Chinese orthography. Huang and Chou (2015) argued that the semantics as the ORL nature of the character writing systems is exactly how it lends itself to serve as the writing system for typologically diversified and phonologically divergent languages including Japanese, Korean, and Vietnamese.

Given that the character writing system is constructed with semantics being its ORL, the crucial issues remains: How do speakers come up with the pronunciations in reading and does the character orthography encode any phonological information at all? The claim that phonological awareness is essential to Chinese reading acquisition has been debated (Huang and Hanley 1995; Taylor 2002; Tzeng and Chen 曾世杰, 陳淑麗 2007). Many studies did not support a relationship between phonological awareness and Chinese reading (Huang and Hanley 1995; McBride-Chang et al. 2005; Shu et al. 2006). Only 30 percent of dyslexic children had a deficit in phonological awareness (Ho et al. 2004). There have been only a few studies dealing with the issue of whether phonological training facilitates character identification (Tzeng and Chen 曾世杰, 陳淑麗 2007; Wang 2017), and if there was a positive effect, the effect size was small (Wang 2017).

One of the most discussed arguments for phonological encoding in the Chinese orthography relies on the fact that the majority of Chinese characters can be analyzed as semantic-phonetic composition (形聲 *xíngshēng*), with the semantic radical and the phonetic part as the two components. It is important to note that the radical is (almost) universal

for all characters, but the phonetic part is not. Based on 2,570 characters from the elementary textbooks in Mainland China, Shu et al. (2003) found that 72 percent of them have a semantic-phonetic composition. Of the 436 characters in the first-grade textbook, only 45 percent belong to the semantic-phonetic composition category. Crucially, only 23 percent of the semantic-phonetic composition characters share the identical phonology (i.e., syllable and tone). This means that only around 16 percent of the characters learned in the first year of school can provide reliable phonological information. This seems to be too low for a learner to rely on for phonological decoding. Although Shu et al. (2000) demonstrated that phonetic parts' regularity does indeed influence character identification performance, the phonological regularity effect is based on the cumulative statistical probability property of the phonetic radical. In other words, it points to the effect of the phonologically regular parts in particular, not the whole system.

Note that there are different approaches to reading training in different Chinese communities. Both Mainland China and Taiwan rely heavily on auxiliary phonological writing systems. The Pinyin romanization is taught in Mainland China and the bpmf NPA in Taiwan. Hence, for subjects in Mainland China and Taiwan, it is currently not possible to differentiate the contribution of the phonological training and the character orthography to their phonological awareness. Interestingly, Hong Kong does not introduce any phonological writing system in reading, and Hong Kong children were shown to be less efficient than Taiwan and Mainland China in their phonological awareness (Huang and Hanley 1995; Cheung et al. 2001; McBride-Chang et al. 2004).

Recall that one of the most critical assumptions about phonological awareness is the role of phonological integrity, i.e., that the use of any linguistic unit in a language L means that the unit must follow the phonology of language L. Interestingly, recent studies on the Mandarin Alphabetic Words (MAW) showed that this basic tenet of linguistic theory might not hold (Huang and Liu 黃居仁, 劉洪超 2017; Ding et al. 2017; Xiang et al. 2020). Note that X-ray is the first attested MAW that was documented more than one hundred years ago, and that the more prominent feature of MAW is the use of Latin (or other non-Chinese) alphabets, in the orthography.

(5) X 光線/射線 [eɪkʰs] /[eɪkʰ̪sʊ̯] *guāngxiàn/shèxiàn* 'X-ray'

What emerged from (5), and especially the widely accepted and quickly increasing MAWs in the last decade or two, is that MAWs somehow licensed violation of Mandarin Chinese phonology. In (5), there are definitely no consonant clusters like 'ks' in Mandarin. And in popular neologisms like Q, 阿 Q, TV, PK, etc., it is clear that Mandarin Chinese simply does not have syllables such as [kwiü], [vi:], or [kei] in its syllabic inventory. Yet, these are not only widely accepted neologisms in Mandarin; it is also widely attested that Mandarin speakers will accept and pronounce them

with a sound that is an approximation of a foreign language (e.g., English) but definitely outside of the Mandarin phonology. The fast increase in MAWs in Mandarin in the last two decades leaves us with two alternatives that are not easily accounted for within the currently accepted paradigm of phonology:

(6) a. that there is no such rule systems as phonology of a language L, nor phonological integrity. Languages have significant freedom in phonological innovation.
b. that language as a complex system does allow the coexistence of more than one phonological system, as long as these incompatible and independent systems are clearly marked.

The alternative solution (6a) is contrary to all existing linguistic theories and would require a totally new paradigm that has not been articulated yet. The alternative solution (6b) would allow several basic theoretical foundations of linguistics to stand, yet requires reinterpretation of the facts that we have discussed so far as orthographic awareness instead of phonological awareness. That is, a set of lexical items adopting the same system of orthographic representation must follow the phonological system encoded by that particular orthography.

Given the theoretical and empirical perspectives introduced earlier, we will now review the reading processing literature on Mandarin Chinese focusing on the issue of phonological or orthographic awareness.

1.4 Linguistic Awareness and Reading Chinese

Theoretically, phonological awareness underlines a speaker's ability to convert continuous speech signals to phonologically significant units. As such, phonological awareness is not dependent on orthography, and we have shown above examples of phonological awareness independent of orthography. The task of phoneme tapping is an excellent example to show phonological awareness without orthography. In this task, a participant is asked to use finger tapping times to indicate the number of sounds for a word. For instance, "cat" as a stimulus expects three times in tapping (Liberman et al. 1974). As such, Castles and Coltheart (2004) argued that phonological awareness was a distal factor instead of a proximal factor for word identification, phonological awareness was not involved in the word identification process, and the opaqueness between orthography and phonology in Chinese would be irrelevant for such tasks. This position is, however, not empirically supported, as most studies in this paradigm, since Liberman et al. (1974), reported some correlation between success rate and intensity of reading training. Hence, the ability to decipher orthography seems to play a role in phonological awareness even when the task itself does not involve orthography.

Another issue related to the potential role of orthography in phonological awareness is the correlation between the linguistic units of reading and the orthographic units. Theoretically, there is a no doubt that word identification is a prerequisite linguistic ability of reading in any language. Chao (1968) pointed out that sociological words in English, i.e., the units identified with spaces in orthography, coincide predominantly with the linguistic words. Hence, identification of units of reading (i.e., linguistic decoding of texts) can take the sociological word as the default units. In addition, an orthographic word in language with phonology as ORL, such as English, can easily analyze it as being composed of smaller phonological units, i.e., 'cat' consists of 'c', 'a', and 't', which in turn represent the phonemes [k], [â], and [t]. Thus, it is quite transparent that phonological awareness is involved in the process of reading a word in a writing system with phonology as ORL.

The sociological words in Chinese, on the other hand, are characters. The most salient difference here is that, unlike English, a Chinese character cannot be analytically decomposed according to phonological rules. As such, identification of an orthographic word in Chinese does not involve explicit application of phonological knowledge as in English. In addition, as mentioned earlier, characters are in fact not words, and identification of words based on characters is not trivial (Huang and Xue 2013). Empirical studies of Chinese based on multi-million-word corpora showed that the average word length in Chinese typically falls in the 1.3 to 1.4 range and that roughly 45 percent of word tokens in Chinese are monosyllabic (Huang et al. 2002). These empirical data in fact invalidate the claim that the correlation between character identification and word identification in Chinese could be as high as 0.8–0.9 (Ho and Bryant 1997; Wang and McBride 2016). In fact, the fallacy of the 0.8–0.9 correlation between characters can be easily debunked by the simple fact of the dominance of disyllabic words in Chinese (over 40 percent of all word types and over 30 percent of all work tokens according to Huang et al. 2002; but certainly exceeding 20 percent in any empirical studies). This erroneous claim that perpetuated the commonly accepted assumption in processing literature that words and characters can be treated as the same psychological construct in fact created an unfortunate fallacy: that a good number of studies on phonological awareness in Mandarin are character-based and hence reported orthographic awareness.

Once this confusion between characters as orthographic units and words as linguistic/phonological units is clarified, many of the puzzles in the literature can be easily interpreted. For instance, Ho and Bryant (1997) observed that rime awareness correlated with character reading but not with Chinese word reading for the first-grader; Wang and McBride (2016) found the same pattern for orthographic awareness. Similarly, the daunting ambiguity of the more than 500 homographs with two more phonological values (i.e., 多音字 *duōyīnzì*), seems to make the basic task of reading a word impossibly complex. However, the phonological value of

these homographs is mostly automatically disambiguated when words are identified. For instance, the monosyllabic word 和 has the pronunciation of *hé* 'and' (or hàn in colloquial Taiwan Mandarin), *hé* in 和平 *hépíng* 'peace', and /he4/ in 附和 *fùhè* 'to echo, to support'. These supposedly ambiguous homographs rarely pose a challenge in a TTS (text-to-speech) system and neither do they pose issues in reading for speakers with lexical knowledge. Hence, the difficulty of dyslexic children in pronouncing the sound of irregular characters (Ho et al. 2006) can be easily explained by their dyslexia, i.e., their difficulties in identifying words affecting the ability to identify the correct pronunciation in lexical contexts.

Another unexpected result in the literature can be similarly accounted for. Surprisingly, statistically significant grade-skipping correlations were reported in several studies of phonological awareness and Chinese character reading. Siok and Fletcher (2001) found that the oddity task was only statistically significantly correlated with character reading in grades 2 and 5, but not in grades 1 and 3. Wei et al. (2014) observed that oddity was statistically significantly correlated with literacy in preschool and second grade but not in first and third grades. A cross-sectional study might yield unstable results because of the varying abilities of the participants for different grades. Yet, the fact that these studies all involve character reading points to potentially a very simple explanation: the disparity between character learning and word learning, a well-known challenge in Chinese curriculum development. Ideally, a Chinese curriculum should pace character learning with word learning; yet, not all the words formed by the same character are of the same level of difficulty, and not all the characters in the same words are of the same difficulty. Thus some grades introduce more new characters, while the next grades may in turn focus on forming new words with these characters and introduce fewer new characters. Such disparity in the pace of character and word reading can easily lead to grade-skipping phonological awareness results in character reading.

Identification of different types of phonological awareness with different tasks: For a language with phonology as the ORL, the relation between a task and the phonological knowledge being tested is fairly transparent. (Landerl and Wimmer 2000; Landerl et al.1997; Paulesu et al. 2001). For instance, both phoneme-based operations and syllable-based operations can be easily designed for English, but syllable-based tasks are much easier to design than phoneme-based tasks due to the transparent status in the orthography. Given that a Chinese character maps to a syllable holistically and does not contain information about phonological segments (Wang 1973), the segment task might not be the right choice. It is well known that the ability to segment phonemes can only be tested after the instruction of phonological writing symbols. Similarly, as the primary instruction topic in the formal instruction for G1 students in Taiwan and Mainland, performance in a blending task is often dependent on the degree of training in the standard educational system. Siok and Fletcher (2001) found that