How Chinese learners decode L2 English words: evidence from a phonics instruction programme

ABSTRACT

The teaching of L2 phonics, defined here as explicit instruction in the symbol-sound correspondences of a second or foreign language, has been enjoying growing popularity in recent years. This is associated with the reemphasis of phonological decoding in different aspects of L2 learning, and learners' difficulty in developing L2 decoding proficiency. However, there has been limited research on L2 phonics instruction, particularly qualitative evidence exploring the nature of learners' decoding errors. For Chinese EFL learners, past studies often stress the negative impact of their morphemic L1 on decoding English words. In contrast, the potential influence of Pinyin knowledge (an alphabetic system for transcribing the spoken Chinese language) has often been overlooked, despite constituting an important aspect of their literacy. This study reports the effects of an English phonics instruction programme, conducted with Chinese university EFL learners, on different categories of English graphemes, classified according to (a) whether they are present or absent in Pinyin, and (b) if present, whether their pronunciations are similar or different across Pinyin and English. In a quasi-experimental design, a group of first-year English majors (n=71) received twelve weeks of English phonics instruction, while a comparison group (n=67) received twelve weeks of English phonology instruction without phonics. The results show that the phonics group made significantly more

progress in decoding some graphemes, especially Pinyin-absent graphemes. Qualitative analysis of participants' decoding identifies three types of common errors in decoding, with the impact of Pinyin being highlighted. Implications for the design of L2 phonics instruction programmes are drawn.

1. INTRODUCTION

1.1 Background

Phonics instruction, defined here as the explicit teaching of the symbol-sound correspondences (SSCs) in an alphabetic writing system, has been the centre of heated discussion in Anglophone countries over the past few decades (Castles et al., 2018). A general consensus nowadays is that phonics instruction plays a positive role in first language (L1) literacy development in alphabetic writing systems (ibid.), although the most effective approaches to teaching phonics and the role of phonics within an overall programme of literacy instruction remain contested (Wyse & Bradbury, 2022). Indeed, systematic phonics instruction has been advocated in official, government-sponsored reviews of early literacy teaching in countries such as Australia (Rowe, 2005), the UK (Rose, 2006) and the US (National Institute of Child Health and Human Development (NICHD), 2000; Ehri & Nunes, 2001).

In recent years, phonics instruction has also become increasingly popular in foreign language (L2) learning contexts. In the UK, Bauckham's (2016) influential modern foreign languages pedagogy review strongly recommends 'direct and systematic

teaching of phonics' in European languages (p.12), a view reinforced in a recent 'Curriculum Research Review' by England's national inspectorate of schools (Ofsted, 2021). Further, in several EFL countries and regions, English phonics now features in the official curricula, including Malaysia and Taiwan (Huo & Wang, 2007). In China, English phonics instruction is currently widely available as an extracurricular programme, but the first author understands from anecdotal evidence that its inclusion in the mainstream school curriculum may be imminent. The most recent Chinese national curriculum has stressed the importance of phonics knowledge in English learning by stating: 'By Year 6 (after four years of English learning), students should develop basic understanding of GPCs [grapheme-phoneme correspondences] and be able to pronounce words using phonics knowledge' (Chinese National Curriculum for English, 2022).

Despite the growth in phonics instruction in L2 settings, research in this area remains limited in both quantity and quality, indicating a need for further work (Li & Woore, 2021). For example, Huo & Wang's (2017) recent review identified 15 studies on the effectiveness of 'phonological-based instruction' (including phonics) for young EFL learners, but only three of these were judged by the authors to be of sufficiently high quality to support causal inferences. Furthermore, many existing L2 phonics instruction programmes use materials originally designed for L1 learners (e.g. Chu & Chen, 2014; Li & Chen, 2016). Such programmes thus fail to take account of learners' existing L1 literacy skills, and may not be optimal for teaching L2 phonics. Careful

error analysis may help develop programmes of instruction which are better tuned to particular learning contexts, including the learners' particular L1-L2 pairing.

To this end, the current paper builds on Li and Woore (2021), which reported on the main effects of a quasi-experimental study of phonics instruction in the context of Chinese EFL learners. Drawing on the same dataset, we now analyse the effects of the teaching in more detail, investigating the accuracy with which participants decoded individual GPCs, and exploring the source of participants' decoding errors. We hope that the results of this analysis will help phonics instruction in this context to be refined in future, targeting those specific GPCs which are most problematic for learners and/or most amenable to instruction.

1.2 The importance of phonological decoding

The popularity of phonics instruction is associated with an emphasis on phonological decoding in different aspects of L1 and L2 learning, where phonological decoding is defined as the ability to generate phonological representations based on orthographic input using knowledge of the language's symbol-sound correspondences. In L1 learning, phonological decoding is generally believed to be an important component of reading comprehension. In the well-established Simple View of Reading (Gough & Tunmer, 1986; Hoover & Gough, 1990), reading comprehension is viewed as the product of decoding and linguistic comprehension. In the early stages of L1 literacy,

knowledge of the language's GPC enables learners to phonologically decode written words which they already know orally, and thus 'discover' their spoken forms and thence their meanings. For instance, by knowing that the English graphemes <c> <a> and <t> (often) correspond to the phonemes /k/ /a/ and /t/, and being able to blend these sounds together in sequence, learners should be able to pronounce the word 'cat' when seeing its written form for the first time, and thereby associate it with its familiar meaning. This serves as the first step towards the development of automatic word recognition and, eventually, efficient reading comprehension (Ehri, 2005).

There is extensive research evidence to show that phonological decoding is also important in L2 reading (Nassaji, 2014). Jeon and Yamashita's (2014) meta-analysis found L2 decoding to be one the strongest correlates of L2 reading comprehension, among L2 vocabulary and grammar (although this does not indicate direction of causality). However, we would argue that the rationale for developing phonological decoding is different for many L2 learners, compared to children acquiring literacy in their L1. L2 learners (at least in instructed, classroom settings) often have limited oral vocabulary knowledge when they are first exposed to the written forms of the language. Hence, in the words of Grabe and Stoller (2011: 36-37), 'one benefit of developing accurate letter-sound correspondences is lost in most L2 settings; that is, L2 students cannot match a sounded-out word to a word that they know orally because they do not yet know the word orally'. Interestingly, in a recent meta-analysis, Odo (2021) found that teaching phonological awareness and/or phonics to

learners of L2 English had a moderate effect on word reading and a large effect on pseudoword reading, but did not explore effects on reading comprehension.

Beyond reading, phonological decoding may still be a foundational skill in various other aspects of L2 learning. For example, Erler & Macaro (2011) found a positive relationship between decoding and motivation, whereby learners with higher L2 decoding proficiency were more likely to choose to continue learning the language in the future (Erler & Macaro, 2011). Other recent research has suggested a relationship between decoding and vocabulary learning (e.g. Hamada & Koda, 2008, 2010), and our previous study (Authors, 2021) provides some initial evidence that this relationship may be causal. In this latter, quasi-experimental study, a group of university students learning English in China received a systematic programme of phonics instruction, covering a wide range of 101 English GPCs, and their outcomes were compared with those of another group who received instruction in English phonology but without explicit phonics. We measured the effects of the phonics instruction on participants' decoding (the accuracy with which they read aloud English words and pseudowords) and their intentional learning of novel English words. The latter was assessed by recall and recognition tests of the new words' phonological and orthographic forms. The results showed that the phonics group outperformed the comparison group in both decoding and word memorization after the instruction, with medium effect sizes.

1.3 The challenges of L2 phonological decoding

Though phonological decoding is believed to underlie various aspects of L2 learning, many L2 learners appear to have difficulty in developing L2 decoding proficiency in the absence of explicit phonics instruction. For instance, studies based in English secondary schools have found that many beginner learners of French make little progress in decoding French words even after years of learning the language, despite the high degree of congruence between the two languages and their writing systems (e.g. Erler, 2003; Erler & Macaro, 2011; Woore, 2009). Similar problems seem to exist even at higher L2 proficiency levels. For instance, Li (2019) found that Chinese university EFL learners performed very poorly on an English word naming task after at least nine years of learning English, showing levels of decoding proficiency equivalent to those of Grade 1 native-speaking children. Similarly, Alghamdi (2020) found deficiencies in L2 decoding skills amongst university students of English in Saudi Arabia: they performed at an equivalent level to Grade 2-3 native speaking children on a word naming task.

The reasons why many learners have trouble developing L2 decoding proficiency are complicated and under-researched. One probable cause is that even experienced L2 learners may encounter difficulty with L2 phonology and phonotactics, which is a foundation for accurate decoding. For instance, Altenberg (2005) found that Spanish learners of English frequently added an epenthetic vowel phoneme before consonant cluster onsets starting with the phoneme /s/ (such as pronouncing *school* as /ɛskul/), in

order to comply with permitted L1 syllable structures. Deterding (2010) found that Chinese learners of English frequently mispronounced voiceless <th> as /s/ (e.g. pronouncing ‘teeth’ as /tiːs/), rather than as the dental fricative /θ/ which does not exist in their L1. These errors can be interpreted as negative cross-linguistic ‘transfer’ (Odlin, 1989) based on L1 phonotactics and phonology.

For learners whose L1 and L2 share the same writing system, such transfer may also be manifested in the automatic triggering of L1 GPCs when decoding L2 words. For instance, Erler (2003), in her study of young learners of L2 French in England, noted that her participants tended to decode French words using English GPCs, such as decoding ‘qui’ (‘who’) as /kwi:/, instead of the correct form /ki:/, even though this is a word they would have encountered frequently and likely knew orally. However, L1 transfer is not always negative in this scenario. For any GPCs shared by both L1 and L2 (e.g. consonant GPC <b> = /b/ in both French and English), L1 phonics knowledge can be beneficial in learners’ L2 decoding.

For learners whose L1 and L2 are linguistically more distant, such as Chinese EFL learners, negative transfer is often assumed. This is due both to issues of phonology and phonotactics, and to a tendency to use processing mechanisms developed in response to their L1 writing system: traditionally, Chinese native speakers are categorized as ‘logographic’ or ‘morphemic’ readers, who process their written L1 at the level of whole characters (Taft et al., 1999). As a result, it is believed that when

decoding English words, Chinese EFL learners are less adept in intraword analysis, leading to poor decoding. Indeed, Hamada and Koda (2008) found that in an English pseudoword decoding task, Chinese EFL learners performed significantly worse than Korean (non-Roman alphabetic) EFL learners, despite having similar English proficiency. The authors attribute this finding to the Korean participants' 'congruent processing experiences' (p. 15) in L1 and L2: both Korean and English are alphabetic writing systems. However, this remains conjecture, because there are many potential confounding variables (the two groups differ in more ways than just their L1 writing systems) and because the lack of any detailed error analysis makes it difficult to ascertain the nature of Chinese EFL learners' poor decoding.

1.4 The decoding of English by Chinese EFL learners: a tale of two systems

Though many believe that the morphemic nature of the Chinese language may be the culprit for Chinese learners' low English decoding proficiency, this may not be the only L1 factor that has an impact. Little research has examined the possible influence of Pinyin on Chinese EFL learners' decoding of English words.

Pinyin is an alphabetic system using the Roman alphabet to denote the pronunciations of Chinese characters. Though sharing the same alphabet, Pinyin and English are very different in terms of orthographic depth: English is generally considered a 'deep' writing system in which the GPCs are often irregular and inconsistent, whereas Pinyin

is phonologically ‘shallow’, with GPCs which are almost always regular and consistent. Moreover, as compared to the very large number of 15,831 syllables in English (Chen, 2015), there are only 413 syllables in Chinese[1], which are made up of 21 onsets (e.g. the /k/ in /ka/) and 38 rimes (e.g. the /a/ in /ka/) (Chen et al., 2016). As a result, native-speaking English children may spend years before becoming proficient in decoding English words, whereas native-speaking Chinese children generally become fluent in Pinyin decoding by the end of Year 1 (Zhang et al., 2020).

Until recently, Pinyin was viewed simply as a tool to facilitate the reading of Chinese characters for native-speaking children, as it is used to ‘transcribe the printed words into phonological codes, which may correspond to concepts they already know orally’ (Lü, 2017: 310). Though Pinyin has been part of the primary school syllabus since 1958 (Li & Rao, 2000), it was not routinely used by adults – until recently. This situation has changed as a result of the prevalence of computers and mobile phones, as Pinyin is the most popular typing method to entre Chinese characters via Western computer keyboards (McBride-Chang, 2012). Figure 1 presents an example of typing the character “我” (which means “I”) with its Pinyin transcription “wo” on a computer.

[1] Syllables are counted here in terms of their constituent phonemes. Chinese also has a system of lexical tones, which greatly expands the number of distinct syllables in the language.

Figure 1. Chinese typing method for the character 我 using Pinyin

Considering that Pinyin has been taught in schools for many years, and furthermore is now actively used by a very large adult population on a daily basis, it seems an omission not to contemplate its possible influence when it comes to research into L1 effects on L2 English learning. Currently, studies in this area have focused mainly on two types of learners: (a) young native-speaking Chinese children who use Pinyin as a tool to learn Chinese characters, and (b) non-native speakers who learn Chinese as a foreign language. A common observation in these studies is that Pinyin knowledge plays a facilitative role in acquiring different aspects of phonological awareness in Chinese, including phonemic awareness, syllabic awareness and tone awareness (Xu & Ren, 2004; Lü, 2017). Moreover, studies involving young Chinese EFL learners suggest that phonological awareness developed through Pinyin has a positive influence on English learning, which is similar to the findings of studies involving other alphabetic L1 speakers learning an alphabetic L2 (e.g. Sparks et al., 2009; Schiff & Calif, 2007). For example, Leong et al. (2005) compared the English phonological awareness of two groups of Chinese children in Years 4 and 5 of primary school, one Beijing group who had previously learned the Pinyin system, and one Hong Kong

group, who (as is routinely the case for Hong Kong children) had not received any instruction in Pinyin. It was found that the Beijing group consistently outperformed the Hong Kong group across all measures of English phonological awareness, including a rhyme detection task, rhyme discrimination task, and phoneme deletion task. Moreover, the Beijing group also performed better in decoding English pseudowords than the Hong Kong group.

The possible influence of Pinyin provides a new angle from which to examine L1 effects on L2 decoding for Chinese EFL learners. Do both written representations of Chinese – the morphemic character system and the alphabetic Pinyin system – come into play when learners decode English words? Are any effects of Pinyin facilitative or disruptive? Furthermore, how does any influence of Pinyin interact with a programme of phonics teaching in L2 English? The effects of phonics instruction on participants' decoding accuracy may differ for individual graphemes, depending on whether or not these graphemes also exist in Pinyin, and if they do, whether their pronunciations in English and Pinyin are similar or different. To answer these questions requires a qualitative analysis of learners' L2 decoding results both before and after a programme of phonics instruction.

1.5 Summary and research questions

To summarize, previous studies have pinpointed the importance of phonological

decoding in different aspects of L2 learning; yet, learners often struggle to develop L2 decoding proficiency in the absence of phonics instruction. Research has begun to indicate that this situation can be improved through systematic teaching of L2 phonics, although further research is required to confirm this. Currently, the teaching of L2 phonics is gaining popularity in different educational contexts, especially in EFL countries. However, many existing L2 phonics instruction programmes use materials designed for L1 learners; given the differences described above between L1 and L2 contexts, it is conceivable that these materials may not be optimal for teaching L2 phonics, and further research is needed to explore this issue. In particular, a more detailed understanding is required of which particular GPC are problematic for particular populations of learners, and whether some GPC may show greater improvement than others as a result of explicit phonics instruction.

For Chinese EFL learners, though past research has stressed the negative influence of the character-based Chinese writing system on decoding English words, it remains unclear to what extent Pinyin also plays a role in the process. To facilitate the understanding of Chinese learners' English decoding, a nuanced qualitative analysis is required of their pronunciations of English words, and how these change as a result of phonics teaching. This may benefit the design of future phonics instruction programmes.

To help understand the above issues, the current study analyses the errors made by

Chinese university EFL students in their phonological decoding of English, both before and after a programme of phonics instruction, with a particular focus on the possible impact of Pinyin knowledge. It also examines the extent to which participants' decoding of various English graphemes can be improved through a programme of phonics instruction. By zooming in on the pronunciation of individual graphemes, this study complements our earlier paper (Li & Woore, 2021), in which we report on the effects of the same phonics programme in terms of the overall accuracy of participants' English decoding and the effects on their intentional word learning. Two research questions are addressed:

(1) To what extent does phonics instruction lead to improved decoding of different categories of L2 English graphemes, classified according to their relationship with Pinyin graphemes?

(2) What are the common errors in participants' decoding of English graphemes, and to what extent can these errors be attributed to the influence of Pinyin?

2. METHOD

2.1 Participants

Participants were 138 first-year English majors studying in four intact classes in two Chinese universities, as reported in Li and Woore (2021). They were all Chinese native speakers who had been in Chinese-medium instruction before entering university. One class in each university was randomly assigned to the intervention (phonics) group, and the other to the comparison (phonology) group. Before

participating in the current study, participants had had approximately nine years of English learning experience, focusing mostly on vocabulary and grammar. Speaking and listening generally received less attention, as these skills were not regularly assessed (Zhang, 2016). Participants had not received any explicit phonics instruction prior to this study.

Prior to the intervention, some baseline information was gathered, namely participants' scores in the National College Entrance English Exam (NCEEE), an English proficiency measure, and the British Picture Vocabulary Scale (BPVS). It was found that the intervention group (N = 71) and the comparison group (N = 67) were matched on both measures, as revealed by two one-way ANOVA (NCEEE: $F(1, 136) = 1.76$, $p = .19$; BPVS: $F(1, 136) = .03$, $p = .86$).

2.2 Intervention

A quasi-experimental pre-/post-test design was adopted. The intervention group received 12 weeks of systematic instruction in English phonics, while the comparison group received 12 weeks of systematic instruction in English phonology, as originally planned in their syllabus.

The phonics instruction consisted of two parts: (a) instruction in 101 English GPCs and 27 common word endings (e.g. *–tion*, *-al*) and (b) the sounding-out of structured

reading materials to practise the newly-taught GPC knowledge. A detailed description of the phonics instruction, including contents and activities, can be found in Li and Woore (2021).

It is worth pointing out that, though the phonics instruction presented a relatively comprehensive picture of English GPCs, it clearly did not, and could not possibly, cover all the probabilities of GPCs in the language, as English is an orthographically deep language (Geva & Wang, 2001). This is especially true for the vowel graphemes, as they may correspond with multiple phonemes depending on the orthographic context. As a result, only GPCs that fit the following two criteria were instructed. First, when the GPCs have well-established, canonical rules, these rules were taught to the participants (e.g. grapheme <oi> is pronounced /ɒɪ/). Secondly, when a grapheme can be decoded in several different ways, only its most common pronunciation was taught to the participants, as recorded in Gontijo *et a*l. (2003): e.g. <ou> → /aʊ/. In other words, some simplification of the true picture was accepted in the phonics instruction programme, so that the most important GPC rules could be taught more easily.

The phonology instruction received by the comparison group consisted of three parts: (a) basic concepts in phonology, such as syllable, stress and rhythm; (b) systematic instruction in the 44 English phonemes, with emphasis on the pronunciation of these phonemes and distinguishing between similar-sounding phonemes; and (c)

pronunciation skills, such as linking, intonation and stress patterns. No explicit instruction in English GPCs was provided in the phonology instruction programme. We acknowledge that the comparison group might have had opportunities for incidental learning of English GPCs, as both written and spoken forms of words were provided when exemplifying a given phoneme. However, the teacher did not draw attention to the English GPC. Thus, the key difference between the two groups was the presence or absence of explicit phonics instruction.

2.3 Data collection

Participants' English decoding was measured before and after the intervention – henceforth time 1 (t1) and time 2 (t2) – using a subtest of the Woodcock Reading Mastery Test (Woodcock, 2011). At each testing point, participants completed one of two forms of the subtest (A and B: see Table 1), requiring them to read 28 pseudowords of increasing difficulty. The reason why pseudowords were used in the test was to make sure that participants did not have any prior knowledge of the test items, and so could not rely on pre-stored pronunciations of the words as wholes: thus, they were forced to process the items at a sub-lexical level, using their knowledge of English GPC.

A pilot test was conducted with 20 participants who were randomly drawn from the two participating universities and who did not participate in the main study. Each pilot

participant was asked to read aloud both Form A and Form B of the decoding test, and their pronunciation of each pseudoword was scored as either correct or incorrect. No significant difference was found between the mean scores achieved by pilot participants in the two test forms, as revealed by a t-test (Form A: mean = 10.3, *SD* = 1.6; Form B: mean = 10.8, *SD* = 1.4; $t(29) = -.77$, $p = .45$). This suggests that the two test forms were of equivalent difficulty for this population.

Table 1. Stimuli in the Phonological Decoding Test (from Woodcock, 2011)

Form A	bab, op, dee, bim, tay, yee, pog, shum, plip, dud's, whie, bufty, vunhip, knaf, twem, adjex, yeng, laip, zirdn't, straced, cedge, wrey, whumb, knoink, bafmotbem, monglustamer, pnir, ceisminadolt
Form B	bab, op, ree, raff, dat, glack, hend, weaf, chur, tayed, ful's, rejune, weat, sess, depine, wrault, throbe, gouch, brecked, darlanker, cigbet, mancingful, squow, cyr, quiles, untroikest, pelnidlum, byrcal

In the main study, a split-block counterbalanced design was used for the decoding test. The participants were randomly assigned to read one of the two test forms at t1 and the other at t2. The test was carried out individually in a quiet classroom. Before the test began, the participants were told specifically that it is natural that they would not know any of these words, as they are pseudowords, and they should just read them as they thought the words should be pronounced. This was to avoid demoralizing the participants and to obtain as complete a dataset as possible.

The participants were asked to read through the word list at their own pace. This was different from the method in many previous studies, where the test items appeared

only for a fixed and usually short period of time (e.g. Hamada & Koda, 2008, 2010). However, some researchers have pointed out the importance for learners of employing conscious processing in order to overcome the automatic triggering of L1 symbol-sound associations when decoding L2 words (e.g. Woore, 2021). This is echoed in Erler (2003), where the English learners of French who took the longest time to decode the written French stimuli actually pronounced them most accurately – it being assumed that this was because they took their time to think about how to decode them accurately in French, rather than simply reading them automatically as English words. As a result, a time limit was not imposed in the current study, lest it put pressure on participants to pronounce words quickly and thus encourage the use of automatic L1 processing.

2.4 Data processing

Participants' pronunciations at both time points were transcribed using the International Phonetic Alphabet (International Phonetic Association (IPA), 2015). All the GPCs appearing in the decoding test were taught in the phonics instruction programme (except for the GPC <pn> = /n/ in Form A, which was thus excluded from the analysis). Graphemes that correspond to regular English spellings of vowels (henceforth 'vowel GPCs') and those that correspond to regular English spellings of consonants (henceforth 'consonant GPCs') were analysed separately. This is not only because phonemes are conventionally divided into these two articulatory categories (e.g. IPA, 2015), but also because participants seemed to encounter more difficulty in

decoding vowels than consonants, based on the first author's observations in the pilot study. Where a grapheme has both consonantal and vocalic phonemic realisations (as in the case of <y>, which could be pronounced as both /j/ as in *yam* and /aɪ/ as in *my*), these cases were analysed separately. Likewise, when a single grapheme has two vocalic or two consonantal phonemic realisations (as in the case of <c> = /s/ or /k/), the two GPCs were analysed separately.

As reviewed in the previous section, a crucial goal of the current study was to examine whether Pinyin GPCs have any impact on participants' decoding. Taking this into consideration, the graphemes in the test were also categorized based on their relationship to Pinyin GPCs. Drawing on the evidence of previous studies involving learners whose L1 and L2 share the same alphabet (e.g. Erler, 2003), two factors that potentially impact participants' decoding were hypothesized, namely: (1) whether an English grapheme also exists in Pinyin; and if yes, (2) whether its corresponding phoneme is similar or different in Chinese.

Based on these hypotheses, the consonant graphemes in our decoding test were divided into three categories:

(a) Pinyin-congruent consonant graphemes: namely English consonant graphemes that also exist in Pinyin and map onto the same (or closely similar) phonemes in Chinese (19 out of the 36 consonant GPCs in the test).

(b) Pinyin-incongruent consonant graphemes: namely English consonant graphemes that also exist in Pinyin but have different corresponding phonemes in Chinese (4 out of 36 consonant GPCs in the test).

(c) Pinyin-absent consonant graphemes: namely English consonant graphemes that do not exist in Pinyin (13 out of 36 consonant GPCs in the test).

The vowel graphemes in the test were also divided into three categories. Given that some English vowel graphemes have more than one possible realization (e.g. <a> is pronounced differently in *cat* and *ago*), this was taken into account in the categorization. Note that there are no vowel graphemes in Pinyin which have the same pronunciation as the English graphemes covered in the phonics instruction; hence, there is no category of 'Pinyin-congruent vowel graphemes'. The three categories of vowel graphemes were therefore as follows:

(d) Pinyin-incongruent vowel graphemes with a single realization: namely those English vowel graphemes that also exist in Pinyin, whose English realization is different from the Pinyin one, and which are associated with only one English phoneme in the phonics instruction programme (4 out of 28 vowel GPCs in the test).

(e) Pinyin-incongruent vowel graphemes with multiple realizations: namely those vowel graphemes that also exist in Pinyin, whose English realizations are different from the Pinyin one, and which are associated with more than one English

phoneme in the phonics instruction programme (9 out of 28 vowel GPCs in the test).

(f) Pinyin-absent vowel graphemes: namely those vowel graphemes that do not exist in Pinyin, and which are associated with only one corresponding phoneme in the phonics instruction programme (15 out of 28 vowel GPCs in the test).

The six categories of GPCs are summarized in Table 2. For ease of comparison, Pinyin pronunciations are also provided for the Pinyin-incongruent GPCs.

Table 2. Six categories of GPCs

Category	English GPC	Corresponding Pinyin GPC (where different from English)
A: Pinyin-congruent consonants	<p>=/p/(op) , <k>=/k/(darlanker), <h>=/h/(hend), <d> = /d/(dud's), <b>=/b/(bab), <sh>=/ʃ/(shum), <f>=/f/(bufty), <l>=/l/(laip), <t>=/t/(tay), <j>=/dʒ/(rejune), <m>=/m/(shum), <z>=/z/(zirdn't), <r>=/r/(throbe), <ch>=/tʃ/(chur), <w>=/w/(weaf), <g>=/g/(gouch), <n>=/n/(depine), <s>=/s/(straced) , <y>=/j/(yee)	
B: Pinyin-incongruent consonants	<s>=/z/(ceisminadolt), <c>=/s/(cigbet), <c>=/k/(byrcal), <x>=/ks/(adjex)	<s>=/s/, <c>=/ts^h/, <x>=/ɕ/
C: Pinyin-absent consonants	<ed>=/d/(tayed), <ss>=/s/(sess), <ff>=/f/(raff), <wh>=/w/(whie), <ck>=/k/(brecked), <wr>=/r/(wrault), <kn>=/n/(knoink), <ng>=/ŋ/(yeng), <mb>=/m/(whumb), <dge>=/dʒ/(cedge), <v>=/v/(vunhip), <th>=/θ/(throbe), <qu>=/kw/(quiles)	
D: Pinyin-incongruent vowels (single realization)	<ie>=/aɪ/(whie), <ai>=/eɪ/(laip), <ei>=/iː/(ceisminadolt), <ou>=/aʊ/(gouch)	<ie>=/je/,<ai>=/ai/, <ei>=/ei/, <ou>=/ou/

E: Pinyin-incongruent vowels (multiple realizations)	<e>=/ə/(bafmotbem), <e>=/iː/(untroikest), <e>=/e/(cedge), <a>=/ə/(byrcal), <a>=/æ/(glack), <u>=/ə/(pelnidlum), <u>=/ʌ/(vunhip), <o>=/ɒ/(bafmotbem), <i>=/ɪ/[2](cigbet)	<e>=/ɤ/, <a>=/ɑː/, <u>=/u/, <o>=/o/, <i>=/i/
F: Pinyin-absent vowels	<a-e>=/eɪ/(straced), <u-e>=/uː/(rejune), <i-e>=/aɪ/(depine), <ir>=/ɜː/(zirdn't), <ey>=/eɪ/(wrey), <au>=/ɔː/(wrault), <ar>=/ɑː/(darlanker), <o-e>=/əʊ/(throbe), <ow>=/əʊ/(squow), <ay>=/eɪ/(tay), <oi>=/ɒɪ/(untroikest), <er>=/ə/(darlanker), <ea>=/iː/(weaf), <ee>=/iː/(ree), <y>=/iː/(yee)	

As the participants were non-native speakers of English, some degree of Chinese colouring in their decoding was considered inevitable and thus permitted. For instance, decoding the English grapheme <r> as /ʐ̩/, which is how it is decoded in Pinyin, was deemed acceptable, as it sounds similar to the English phoneme /ɹ/ and there is little likelihood that other English phonemes can be confused with /ʐ̩/. Similarly, pronouncing the English grapheme <sh> with Chinese colouring as /ʂ/ was also considered acceptable. However, decoding the grapheme <th> as /s/ was deemed unacceptable: though /s/ is the Chinese phoneme that arguably most closely resembles the English phoneme /θ/, there are other English graphemes that should be decoded as /s/, and failing to make a distinction between these two phonemes can result in poor comprehension and wrong representations of words containing them (e.g. a failure to distinguish between the pronunciations of 'sing' and 'thing').

[2] In the case of <o> and <i>, these have additional realizations which were included in the phonics instruction (<o> = /ə/ as in Lond**o**n; <i> =/aɪ/ as in ch**i**ld), but these did not appear in the decoding test and so are not included in this table.

Each GPC in the decoding test scored one point if it was decoded correctly, otherwise zero points. Twenty percent of the results were randomly selected and re-scored by both the first author and a colleague, who was fluent in both Mandarin and English and had a Master's degree in Applied Linguistics. The intra-rater agreement was 94.3% and inter-rater agreement was 90.5%. Participants' scores in each GPC category as listed above were calculated. Given that the number of occurrences in each GPC category was different, the scores were converted into accuracy percentages for ease of comparison.

3. RESULTS

3.1 RQ1: The effects of phonics instruction on different categories of English graphemes

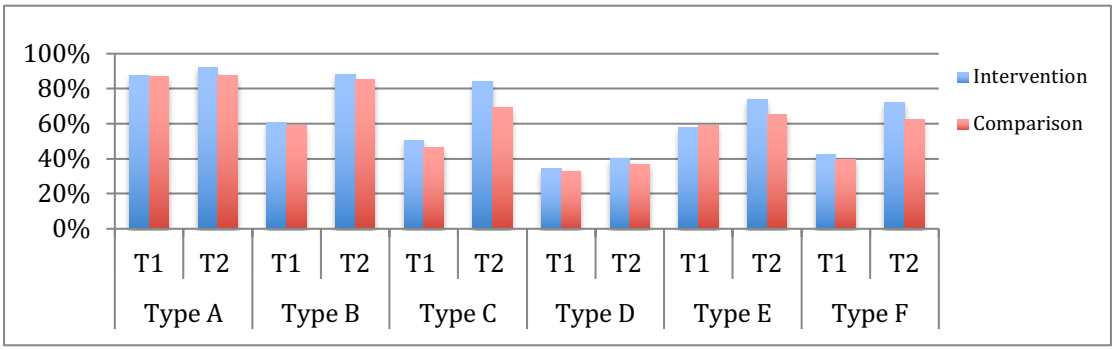
The mean accuracy of the six types of graphemes before and after the instruction programmes is shown in Table 3 and Figure 2.

Table 3. Mean decoding accuracy of different categories of English GPCs

		T1 mean	*T1 SD*	*T2 mean*	*T2 SD*
Type A: Pinyin-congruent consonants	*Intervention*	87.6%	7.5%	92.1%	3.5%
	Comparison	87.1%	7.6%	87.6%	14.3%
Type B: Pinyin-incongruent consonants	*Intervention*	*61.0%*	8.8%	*88.6%*	*11.9%*
	Comparison	*59.2%*	10.3%	*85.5%*	*20.8%*
Type C: Pinyin-absent consonants	*Intervention*	50.9%	18.0%	84.4%	9.7%
	Comparison	46.6%	16.8%	69.6%	18.2%
Type D: Pinyin-incongruent	*Intervention*	34.7%	16.1%	40.3%	11.8%

vowels (single realization)	*Comparison*	33.3%	21.1%	37.4%	20.8%
Type E: Pinyin-incongruent vowels (multiple realizations)	*Intervention*	58.0%	14.5%	73.9%	11.6%
	Comparison	59.2%	12.0%	65.4%	17.3%
Type F: Pinyin-absent vowels	*Intervention*	42.8%	11.0%	72.4%	12.3%
	Comparison	40.1%	9.7%	62.6%	10.6%

Figure 2. *Mean decoding accuracy of different categories of English GPCs*



To examine whether phonics instruction led to improved decoding of each different category of English graphemes, a series of mixed factorial ANOVA with one within-subjects variable (time) and one between-subjects variable (condition) was conducted. A Bonferroni correction for multiple comparisons was made, adjusting the alpha level for statistical significance to .008. The results are as follows.

For type A graphemes (Pinyin congruent consonants), there was no significant main effect of time, $F(1, 136) = 5.50$, $p = .02$, and no significant main effect of condition, $F(1, 136) = 4.85$, $p = .03$. The interaction between time and condition was not

significant, $F(1, 136) = 3.62$, $p = .06$.

For type B graphemes (Pinyin-incongruent consonants), there was a significant main effect of time, $F(1, 136) = 471.14$, $p < .001$. There was no significant main effect of condition, $F(1, 136) = 2.37$, $p = .13$. The interaction between time and condition was also non-significant, $F(1, 136) = .13$, $p = .72$.

For type C graphemes (Pinyin-absent consonants), there was a significant main effect of time, $F(1, 136) = 615.3$, $p < .001$, and a significant main effect of condition, $F(1, 136) = 10.8$, $p = .003$. The interaction between time and condition was also significant, $F(1, 136) = 14.61$, $p < .001$.

For type D graphemes (Pinyin-incongruent vowels with a single realization), there was a significant main effect of time, $F(1, 136) = 11.74$, $p < .001$. There was no significant main effect of condition, $F(1, 136) = .09$, $p = .76$. The interaction between time and condition was also non-significant, $F(1, 136) = .07$, $p = .93$.

For type E graphemes (Pinyin-incongruent vowels with multiple realizations), there was a significant main effect of time, $F(1, 136) = 69.65$, $p < .001$. There was no significant main effect of condition, $F(1, 136) = 3.39$, $p = .07$. The interaction between time and condition was significant, $F(1, 136) = 13.42$, $p < .001$.

For type F graphemes (Pinyin-absent vowels), there was a significant main effect of time, $F(1, 136) = 623.40$, $p < .001$, and a significant main effect of condition, $F(1, 136) = 12.13$, $p = .001$. The interaction between time and condition was also significant, $F(1, 136) = 15.42$, $p < .001$.

The above results suggest that the phonics instruction was effective (relative to a comparison group receiving no phonics) in promoting the decoding of Pinyin-absent consonant graphemes, Pinyin-incongruent vowel graphemes with multiple realizations and Pinyin-absent vowel graphemes. The phonics instruction did not lead to improved decoding of Pinyin-congruent consonant graphemes, Pinyin-incongruent consonant graphemes or Pinyin-incongruent vowel graphemes with a single realization.

3.2 RQ2: Frequent errors in decoding different categories of English graphemes

For type A graphemes (*Pinyin-congruent consonants*), the most frequent error – both before and after the intervention – was the omission of graphemes when decoding a consonant cluster. This was mostly observed in long words with multiple syllables, such as *monglustamer* and *untroikest*. The omission errors that were observed are listed in Table 4, together with the number of times each error occurred in the particular word shown, broken down by experimental group and time point. For example, at time 1, four out of the 71 intervention group participants omitted the

second <b> in *bafmotbem*, as did six out of the 69 comparison group participants.

Table 4. Omission of type A graphemes (Pinyin-congruent consonants)

Omitted grapheme (and its correct pronunciation)	Target word (the omitted consonant is underlined)	Illustrative examples of errors	Number of Errors			
			Intervention group (n=71)		Comparison	
			t1	t2	t1	t2
<b> (/b/)	*baf*mot*b*em	/bæfmɒtem/, /bæfmɒtəm/	4	3	6	4
<t> (/t/)	*baf*mo*t*bem	/bæfmɒbem/, /bæfməubəm/	2	1	11	3
<f> (/f/)	ba*f*motbem	/bæmɒtbem/, /bæməutbəm/	6	0	8	2
<g> (/g/)	mon*g*lustamer	/mɒnlʌstəmə/, /mɒnlju:stəmə/	30	1	31	10
	ci*g*bet	/sɪbet/, /sɪbɪt/	13	0	11	1
<s> (/s/)	monglu*s*tamer	/mɒnglʌtəmə/, /mɒnglju:təmə/	22	0	19	6
	untroike*s*t	/ʌntrɒɪkət/, /ʌntrɒɪkɪt/	20	1	25	10
<l> (/l/)	pelnid*l*um	/pelnɪdəm/	5	0	6	4

For type B graphemes (Pinyin-incongruent consonants), the most frequent error at both time points was Pinyin-resembling approximation, defined as decoding a

grapheme as its Pinyin pronunciation (or similar-sounding pronunciation). These errors are shown in Table 5.

Table 5. Approximation in type B graphemes (Pinyin-incongruent consonants)

Mispronounced grapheme (and its correct pronunciation)	Target word	Error	Number of Errors			
			Intervention		Comparison	
			t1	t2	t1	t2
<x> (/ks/)	*adjex*	/adjes/	32	23	36	28
<s> (/z/)[3]	*ceisminadolt*	/siːsminədəʊlt/, /saɪsminədəʊlt/	9	3	9	6

For type C graphemes (Pinyin-absent consonants), both vowel epenthesis and Pinyin-resembling approximation were observed (Tables 6 and 7).

[3] The scoring manual for the Woodcock Reading Mastery Test (Woodcock, 2011) gives the correct pronunciation of <s> here as /z/. We have therefore interpreted this error as a 'pinyin-resembling approximation'. However, we do so with caution: we are aware that <s> is often pronounced as /s/ in English and indeed the second author (a native speaker of English) would pronounce it as /s/ in this pseudoword.

Table 6. Epenthesis in type C graphemes (Pinyin-absent consonants)

Mispronounced Grapheme (and its correct pronunciation)	Target word	Illustrative example of errors	Number of Errors			
			Intervention		Comparison	
			t1	t2	t1	t2
<kn> (/n/)	*knaf*	/kənaf/	28	3	31	21
		/kenaf/	0	0	2	0
	knoink	/kənaf/	30	5	28	24
		/kinaf/	5	0	5	7
<wr> (/r/)	*wrault*	/wərɔːlt/	8	0	11	6
	wrey	/wereɪ/	6	0	4	2
		/wəreɪ/	3	0	4	1
<wh> (/w/)	*whumb*	/wəhʌm/	5	0	4	0
		/wiːhʌm/	3	0	0	0
<dge> (/dʒ/)	*cedge*	/sediːdʒ/	28	6	20	17
		/sededʒ/	23	6	19	23
<qu> (/kw/)	*squow*	/skjuːrəʊ /	33	8	28	21
		/skɔːrəʊ /	8	3	11	13
	quiles	/kjuːləs/	31	10	24	12
		/kwɔːləs/	9	3	14	7

Table 7. Approximation in type C graphemes (Pinyin-absent consonants)

Mispronounced Grapheme	Target word	Illustrative example of errors	Number of Errors			
			Intervention		Comparison	
			t1	t2	t1	t2
<v> (/v/)	*vunhip*	/wʌnhɪp/	11	1	9	2
<th> (/θ/)	*throbe*	/srəʊb/	47	39	46	43

For type D graphemes (Pinyin-incongruent vowels with a single realization), the most common error was Pinyin-resembling approximation, as summarized in Table 8.

Table 8. Approximation in type D graphemes (Pinyin-incongruent vowels with a single realization)

Mispronounced grapheme (and its correct pronunciation)	Target word	Illustrative example of errors	Number of Errors			
			Intervention		Comparison	
			t1	t2	t1	t2
<ei> (/iː/)[4]	*ceisminadolt*	/ceɪzmɪnədəʊlt/	30	20	32	28
<ie> (/aɪ/)	*whie*	/wɪe/	26	25	24	21
<ou> (/aʊ/)	*gouch*	/goutʃ/	50	41	50	46
<ai> (/eɪ/)	*laip*	/laip/	48	41	45	43

For type E graphemes (Pinyin-incongruent vowels with multiple realizations), various kinds of errors were observed (Table 9), though Pinyin approximation is still very common (examples being highlighted in bold in the 'error' column).

[4] We are aware that <ei> can be pronounced /eɪ/ in English, as in *beige* and *reign*. However, according to Gontijo et al. (2003), this pronunciation is rare, occurring in only 1% of instances of <ei>. Therefore, we believe it is appropriate to classify this pronunciation as a Pinyin-resembling approximation error.

Table 9. Approximation and other errors in type E graphemes (Pinyin-incongruent vowels with multiple realizations)

Mispronounced grapheme (and correct pronunciation)	Target word	Error (Pinyin approximation in **bold**)	Number of Errors			
			Intervention		Comparison	
			t1	t2	t1	t2
<a> (/æ/)	*bab, knaf, bafmotbem, raff, dat, glack, darlanker, mancingful, adjex,*	**/a:/**	14	5	13	10
		/e/	1	3	2	4
		/eɪ/	5	3	5	6
<a> (/ə/)	*byrcal, monglustamer, ceisminadolt*	**/a:/**	21	6	21	17
		/æ/	11	8	9	10
		/eɪ/	11	19	10	10
<e> (/e/)	*twem, adjex, yeng, cedge, hend, sess, brecked, pelnidlum*	/iː/	20	6	22	17
		/æ/	18	14	22	11
<e> (/iː/)	*rejune, depine*	/e/	28	14	28	28
<e> (/ə/)	*untroikest*	/e/	18	13	18	9
		/iː/	14	11	14	12
<u> (/ʌ/)	*shum, dud's, bufty, vunhip, whumb, ful's, untroikest*	**/u/**	21	14	25	21
		/ɣ/	13	8	7	8
		/æ/	13	12	7	8
<u> (/ə/)	*mancingful, pelnidlum*	**/u/**	39	23	43	30
		/ʌ/	11	19	4	14
<o> (/ɒ/)	*op, monglustamer*	/æ/	21	15	14	19
<o> (/əu/)	*ceisminadolt*	**/ɔ/**	28	10	32	25
<i> (/ɪ/)	*bim, plip, vunhip, cigbet, pelnidlum*	/aɪ/	7	4	6	10

For type F graphemes (Pinyin-absent vowels), various kinds of errors were observed, and the pattern was less clear. However, a frequent error in decoding so-called 'split digraphs' in this category was the epenthesis of an additional vowel at the end of the word, such as decoding *rejune* as /rɪdʒuːnə/ (Table 10).

Table 10. Epenthesis in type F graphemes (Pinyin-absent vowels)

Mispronounced grapheme (and correct pronunciation)	Target word	Illustrative example of errors	Number of Errors			
			Intervention		Comparison	
			t1	t2	t1	t2
<u-e> (/uː/)	*rejune*	/rɪdʒuːnə/, /redʒuːnə/	17	1	19	17
		/redʒuniː/	13	0	9	1
<i-e> (/aɪ/)	*depine*	/dɪpaɪnə/, /dɪpɪnə/, /depɪnə/	23	3	21	14
		/dɪpɪniː/, /depɪniː/	7	0	6	1
	quiles	/kwɒləs/, /kwɒləz/, /kwaɪləs/, /kɪləs/, /kwɪləs/, /kjuːləs/	26	3	16	12
		/kwɒliːs/, /kwɒliːz/, /kɪliːs/, /kwɪliːs/, /kjuliːs/	14	3	20	16
<o-e> (/əʊ/)	*throbe*	/θrəʊbə/, /srəʊbə/	4	0	2	0
		/θrəʊbiː/, /srəʊbiː/, /θrɒbiː/ /srɒbiː/	18	1	14	10
<a-e> (/eɪ/)	*straced*	/stræsəd/, /stresəd/, /strʌsəd/	8	0	6	2
		/stræsiːd/ /stresiːd/ /strʌsiːd/	9	6	11	12

In addition, some errors were found in the analysis which were not easily classifiable in terms of individual graphemes. One such type was whole-word errors, when participants wrongly decoded a pseudoword as a real word with similar orthography (Table 11).

Table 11. Whole-word errors

Target word	Illustrative examples of errors	Number of Errors			
		Intervention		Comparison	
		t1	t2	t1	t2
bab	bad	7	1	9	5
whie	well, wine, while, where, wheel	25	5	21	15
twem	twin, team	10	2	12	13
yeng	year, young	15	0	13	6
wrey	worry, when, well, wine, white, wall, way, very	30	5	35	26
knoink	knock, kick	29	14	26	17
wrault	worried, result, wallet, wreck	12	3	15	10
dat	date	15	6	19	10
cyr	cry	20	0	17	12
quiles	quiz	0	0	3	3
weat	water, wake	0	0	5	0

Finally, another type of error was misordered graphemes, where graphemes were decoded in the wrong order, but without thus forming a real word (Table 12).

Table 12. Misordered graphemes

Target word	Illustrative examples of errors[5]	Number of Errors			
		Intervention		Comparison	
		t1	t2	t1	t2
knoick	knocki (/nɒki:/, /kənɒki:/)	10	0	9	4
bufty	bfuty (/bfjuti:/, /bəfuti:/)	8	2	11	8
mancingful	mangcinful (/mangsɪnfəl/)	5	0	3	3
byrcal	brycal (/bri:kəl/, /braɪkəl/)	10	0	8	4
pelnidlum	pelindlum (/pelɪndləm/)	5	0	3	3

[5] The spellings in this column are the 'reconstruction' of the written forms of participants' decoding

4. DISCUSSION

4.1 The effects of phonics instruction on different categories of English graphemes

Addressing RQ1, the results suggest that the phonics instruction was effective in promoting the decoding of some types of graphemes: Pinyin-absent consonant graphemes, Pinyin-incongruent vowel graphemes with multiple realizations and Pinyin-absent vowel graphemes. However, the phonics instruction did not lead to significantly more progress in decoding Pinyin-congruent consonant graphemes, Pinyin-incongruent consonant graphemes or Pinyin-incongruent vowel graphemes with a single realization. The findings suggest that Pinyin plays an active role in participants' decoding and has a clear impact on the results of the phonics instruction. This is evidenced from the following perspectives.

First, though Pinyin-congruent consonant graphemes did not see significantly more progress after the phonics instruction, they already tended to be decoded successfully at t1, with a mean accuracy approaching 90% for both groups. This was expected, as the Contrastive Analysis Hypothesis (Lado, 1957; Ellis, 1994) argues for positive transfer from L1 to L2 when similar properties are shared by the two languages, such as GPCs in this case. However, it is worth noting that the facilitative role of L1 in L2 decoding generally seems to have been documented in learners whose L1 and L2 writing systems are typologically similar. In contrast, a typologically distant L1 has always been argued to hinder the development of L2 decoding proficiency. On this

basis, Koda proposed the Orthographic Distance Effect (2005), arguing that the more similar the L1 and L2 writing systems are, the more easily learners can develop decoding proficiency in L2. Within this framework, Chinese learners are typically characterized as having a logographic (or morphemic) L1 writing system, which is typologically distant from the English writing system. The high accuracy achieved in Pinyin-congruent GPC at t1 suggests that, in L2 writing system research, it is inadequate simply to characterize Chinese L1 learners as 'logographic', as this fails to capture an important aspect of their existing literacy experience, namely their knowledge of the alphabetic Pinyin writing system.

Second, many Pinyin-incongruent graphemes seemed to resist the effects of phonics instruction in our study. This is in line with previous findings involving learners whose L1 and L2 share the same alphabet (e.g. English and French), as the automatic triggering of L1 symbol-sound correspondences may be difficult to combat (Author, 2021). In the current study, it appears that many L1-based associations remained stubbornly entrenched, even after explicit instruction. Among the three types of Pinyin-incongruent graphemes, only Pinyin-incongruent vowels with multiple realizations saw significantly more progress amongst the intervention group between t1 and t2. One possible explanation is that these graphemes are very frequently encountered – more frequently than the other two types of Pinyin-incongruent GPCs (Gontijo et al., 2003) – thus giving learners more opportunities to practise them. Overall, it may be that more time and practice are needed – beyond the 12 weeks of

the phonics programme in the current study – to help learners combat the automatic application of Pinyin GPCs when decoding Pinyin-incongruent English graphemes. This again suggests that Pinyin plays an important role in the development of English decoding proficiency.

Third, intervention participants made significantly more progress than the comparison group in decoding Pinyin-absent graphemes, both consonants and vowels. On the one hand, this indicates that learners may not learn the pronunciation of such L2 graphemes in the absence of explicit instruction. On the other, it also seems to suggest that teaching L2 graphemes 'from scratch' is somewhat easier than teaching L2 graphemes that already exist in L1, but have different pronunciations. Again, this may be because learners are less likely to be influenced by existing L1-based spelling-sound associations.

Taken together, these findings indicate the need to take Pinyin knowledge into consideration when designing English phonics instruction for Chinese EFL learners. As Pinyin knowledge seems to be facilitative in decoding Pinyin-congruent graphemes, these could be treated more briefly in future phonics programmes. Instead, more time could be allocated to teaching Pinyin-incongruent graphemes and Pinyin-absent graphemes – and to practising the former, in order to help learners to override the automatic triggering of inappropriate Pinyin GPCs when decoding English words.

4.2 Common errors in English decoding

Addressing RQ2, error analysis reveals three frequent types of error in participants' decoding of English, namely omission, epenthesis and approximation. Below, we explore each of these in turn.

First, participants frequently omitted some consonant graphemes at t1, especially in long words with multiple syllables such as *bafmotbem* and *monglustamer*. An interesting observation is that the omitted consonant graphemes are most often found at the start of a consonant string, while the consonant grapheme right before a vowel was almost always pronounced. A possible explanation is that the consonant grapheme right before a vowel was perceived together with the following vowel, which is similar to the onset + rime structure of Chinese syllables. This might also explain why consonant graphemes occurring after a vowel were often omitted, as Chinese syllables always end in a vowel[6]. This observation echoes the findings of some speech analysis studies (e.g. Broselow et al., 1998) that Chinese EFL learners tend to omit syllable codas in their spontaneous speech in English. Thus, an important reason for the omission errors in the current study may not be unfamiliarity with

[6] Some Pinyin words end with the graphemes <n> or <ng>, suggesting a CVC structure (e.g. *man, ping*). However, these are generally interpreted as representing nasalized vowels, giving a phonological structure of CV (Chen, 2016).

particular L2 GPC, but instead, the influence of L1 phonotactics and syllable structure. This might explain why the comparison participants also made fewer errors in this category at t2, as the phonology instruction may have familiarized them with English phonotactics, in turn leading to better decoding of English consonant clusters.

Second, epenthesis was commonly found in the decoding of two kinds of graphemes, namely consonant digraphs/trigraphs (<kn>, <wr>, <qu>, <wh> and <dge>) and split digraphs (e.g. <a – e>). These errors seemed to originate from unfamiliarity with the inventory of English graphemes. For instance, at t1, many participants failed to recognise these graphemes as a single grapheme, and made Pinyin-resembling pronunciations by either adding a vowel grapheme between what they perceived as two separate consonants (e.g. decoding *knaf* as /kənæf/) or by segmenting syllables incorrectly (e.g. decoding *rejune* as /rɪdʒuːnə/). At t2, the intervention participants made markedly fewer epenthesis errors than at t1, while the comparison participants made a similar amount of epenthesis errors. This suggests that, although familiarizing learners with the English phonology system can be beneficial in promoting some aspects of English decoding, it may not be sufficient in itself, especially when it comes to decoding Pinyin-absent and Pinyin-incongruent graphemes.

Third, participants made Pinyin-resembling approximation errors in decoding various kinds of graphemes at t1, mostly Pinyin-incongruent graphemes – more specifically, Pinyin-incongruent consonant graphemes, Pinyin-incongruent vowel graphemes with

a single realization and Pinyin-incongruent vowel graphemes with multiple realizations. This again points to the important role that Pinyin plays in participants' L2 decoding: when participants were not equipped with adequate English GPC knowledge, they tended to fall back on their existing Pinyin GPC knowledge when decoding graphemes shared by Pinyin and English, resulting in Pinyin-resembling pronunciations. This recalls the young beginner-learners of L2 French in Erler (2003), who relied on the 'heuristic of English' to pronounce French words.

At t2, though the error rate of many of these graphemes remained high for both groups, it is worth noting that intervention participants made fewer Pinyin-resembling approximation errors, especially in the decoding of Pinyin-incongruent vowels with multiple realizations. For instance, when decoding the English GPC <a> = /ə/ (as in *ago*), the Pinyin-resembling approximation error (<a> = /a:/) decreased from 29% to 9% for the intervention participants, but other errors (such as <a> = /eɪ/) rose from 15% to 27%. Whilst both /a:/ and /eɪ/ are incorrect pronunciations, the latter may be seen as a 'better' error, insofar as it may indicate that participants had moved beyond the automatic application of the L1 (Pinyin) GPC. This may be seen as a first step towards establishing correct L2 GPCs. It can therefore be treated as progress in the development of L2 decoding, even though the decoded forms are still incorrect. We would observe that this illustrates the value of qualitative analysis in L2 decoding studies, as both pronunciations would simply be treated as wrong in a quantitative analysis.

In addition to the three types of errors discussed above, participants also made some errors beyond individual graphemes, namely whole-word errors and mis-ordering of graphemes. These errors might be seen to indicate a lack of (or at least deficiencies in) intraword analysis, which has been associated with the morphemic nature of the learners' L1 Chinese writing system in previous studies (e.g. Akamatsu, 2003). Though this may be true to some extent, it should be noted that these cases account for only a small percentage of errors in our participants' decoding: of the 56 pseudowords in the two test forms, whole-word errors were found in only 11 of them and mis-ordered graphemes in only 5. Moreover, a further look at these errors suggests that their origin may be more complicated. For instance, the two errors in decoding the word *knoick* were /nɒki:/ and /kənɒki:/, both of which end in a vowel and follow a (CV)CVCV pattern. Similarly, the two errors in decoding the word *bufty* were /bəfjuti:/ and /bəfuti:/, both of which also end in a vowel and follow a CVCVCV pattern. Again, this suggests a potential influence of Chinese phonology, where syllables always consist of an onset and a rime (with no coda). The participants who committed these errors might have been pronouncing these words in a way that fits the pattern of Chinese syllables. Though this remains conjecture, it still suggests that simply attributing participants' errors in English decoding to their morphemic L1 writing system may be an oversimplification.

In summary, error analysis of participants' decoding suggests some possible reasons

behind Chinese EFL learners' difficulties in decoding English words, in turn providing a basis for better understanding the effects of a programme of explicit L2 phonics instruction. In order to help Chinese EFL learners combat these errors, it is of crucial importance to take into consideration their existing Pinyin knowledge, which was often neglected in past research.

4.3 Limitations

Several limitations of the current study must be acknowledged. First, our study was small in scale: whilst we had 138 participants in total, they were drawn from only four intact classes in two Chinese universities. Clearly, we would wish to replicate our study with a larger sample drawn from a wider pool of universities. Second, whilst we checked that the participating classes were matched in terms of English proficiency and vocabulary knowledge, these intact groups may have differed in other ways, of which we were unaware (e.g. motivation for learning English); this may have affected the results of Research Question 1. A larger-scale, adequately powered cluster randomized controlled trial (RCT) would be needed to address this issue (Hutchison & Styles, 2010), which was unfortunately not possible in our case due to resource constraints. Third, the limited nature of our sample also means that our findings cannot be generalized to learners from other backgrounds; more research, involving learners of different ages, proficiency levels and L1-L2 pairings, is needed to explore the optimal way of teaching L2 phonics in different educational contexts. Finally, we

would ideally have included a delayed post-test to examine the extent to which the effects of phonics instruction were maintained over time, but this was regrettably not possible due to scheduling difficulties.

5. CONCLUSIONS

To recapitulate, we adopted a quasi-experimental design to evaluate the effects of L2 phonics instruction, using a sample of Chinese university EFL learners, on the pronunciation of different categories of English graphemes. The main outcomes of the study are reported in Li and Woore (2021); in the current paper, we conducted detailed analysis of participants' decoding, both before and after the phonics instruction, in order to identify the origin of their decoding errors, with particular attention being paid to the potential influence of Pinyin GPCs. The results showed evidence of a strong influence of Pinyin on participants' phonological decoding of L2 English.

The phonics instruction was effective (relative to a non-phonics comparison group) in promoting the decoding of Pinyin-absent consonant graphemes, Pinyin-incongruent vowel graphemes with multiple realizations and Pinyin-absent vowel graphemes. However, the phonics instruction did not lead to significantly greater progress in decoding Pinyin-congruent consonant graphemes, Pinyin-incongruent consonant graphemes or Pinyin-incongruent vowel graphemes with a single realization. (There

were no Pinyin-congruent vowel graphemes in our analysis).

Error analysis of participants' decoding mainly revealed three main types of errors, namely grapheme omission, epenthesis and Pinyin-resembling approximation. Some of these errors seemed to arise through the influence of Pinyin or Chinese phonology more broadly – in particular, phonotactic constraints and syllable structure. Whilst some other errors, such as whole-word errors and misordered graphemes, were also observed, which may indicate deficiencies in intraword analysis, these only represented a small proportion of the errors in participants' decoding.

Our findings may be useful in informing future L2 phonics instruction. Though such instruction is increasingly popular in China and many other countries (see section 1.1), how to teach phonics most effectively in L2 needs further examination. The current study tentatively suggests that it is important to take learners' existing literacy skills into account, and to focus on aspects where their L1 and L2 diverge. In the case of Chinese EFL learners, as Pinyin-incongruent graphemes seem particularly challenging, future phonics instruction programmes may target these graphemes specifically and provide plentiful opportunities to practise them, instead of spending equal time on all graphemes.

The findings of this study also point to the need to reconsider the common claim that

Chinese EFL learners can simply be classified as ‘logographic’ or ‘morphemic’ readers for the purposes of their decoding of English words. Given that many Chinese EFL students learned Pinyin at school and now use it on a daily basis for keyboard entry of Chinese characters, the findings of some previous cross-linguistic studies may need to be re-examined in order to take account of this important aspect of participants’ L1 literacy. This applies especially to studies conducted after the new millennium, when mobile phones and computers began to be prevalent. In addition, future crosslinguistic studies may wish to examine further the influence of Pinyin knowledge on various other aspects of Chinese EFL students’ English learning, which remains an under-researched area that could yield fruitful results.

ACKNOWLEDGEMENTS

This research is supported by the Fundamental Research Funds for the Central Universities. The authors wish to thank the three anonymous reviewers for their constructive comments in revising this article. The authors would also like to thank the participating universities, teachers and students.

Sha Li (corresponding author) is a lecturer in the Department of Linguistics, School of International Studies, Zhejiang University, Hangzhou, Zhejiang, China; e-mail sha_li@zju.edu.cn
Robert Woore is Associate Professor in the Department of Education, University of Oxford, Oxford, Oxfordshire, UK; e-mail robert.woore@education.ox.ac.uk

REFERENCES

Akamatsu, N. (2003). The effects of first language orthographic features on second language reading in text. *Language Learning*, 53 (2), 207-231.

Alghamdi, H. (2020). L2 decoding and reading comprehension in Saudi foundation year ESL students: an investigation into readers' decoding abilities, outputs, experiences, and processes, and the contribution of their L2 decoding proficiency to their L2 reading comprehension. Ph.D thesis. University of Oxford.

Altenberg, E. P. (2005). The perception of word boundaries in a second language. *Second Language Research*, *21*, 325-358.

Bauckham, I. (2016). Modern foreign languages pedagogy review. A review of modern foreignlanguages teaching practice in key stage 3 and key stage 4. Teaching Schools Council: London.

Broselow, E., Chen, S. I., & Wang, C. (1998). The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition*, *20*, 261-280.

Campbell, S. (2020). Teaching phonics without teaching phonics: Early childhood teachers' reported beliefs and practices. *Journal of Early Childhood Literacy*, *20*, 783-814.

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*, 5-51.

Chinese National Curriculum for English. (2022). http://www.moe.gov.cn/srcsite/A26/s8001/202204/W020220420582349487953.pdf

Chen, L. L. (2016). Hanyu pinyin. In C. Sin-Wai (ed.). *The Routledge encyclopedia of the Chinese language* (pp. 522-542). London: Routledge.

Chen, M. Y. (2015). Discussion 1: A paradox: Merger or Compounding— which comes first? comments on Geoffrey Sampson's article. *Journal of Chinese Linguistics*, *43*, 692-696.

Chen, N. F., Wee, D., Tong, R., Ma, B., & Li, H. (2016). Large-scale characterization of non-native Mandarin Chinese spoken by speakers of European origin: Analysis on iCALL. *Speech Communication*, *84*, 46-56.

Chu, M. C., & Chen, S. H. (2014). Comparison of the effects of two phonics training programs on L2 word reading. *Psychological Reports*, *114*, 272-291.

Deterding, D. (2010). ELF-based Pronunciation Teaching in China. *Chinese Journal of Applied Linguistics*, *33*, 3-15.

Ehri, L. C. (2005) 'Development of Sight Word Reading: Phases and Findings', in M. J. Snowling and C. Hulme (eds.) 2005, The Science of Reading. A Handbook. Oxford: Blackwell, pp. 135-154.

Ehri, L. C. & Nunes, S. R. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Education Research*, *71*, 393-447.

Ellis, R. 1994. The study of second language acquisition. Oxford: Oxford University Press.

Erler, L. (2003). *Reading in a foreign language--- near-beginner adolescents' experience of French in English secondary schools*. Ph.D thesis. University of Oxford.

Erler, L., & Macaro, E. (2011). Decoding ability in French as a foreign language and language learning motivation. *The Modern Language Journal*, *95*, 496-518.

Geva, E., & Wang, M. (2001). The development of basic reading skills in children: A cross-language perspective. *Annual Review of Applied Linguistics*, *21*, 182-204.

Gontijo, P. F., Gontijo, I., & Shillcock, R. (2003). Grapheme-phoneme probabilities in British English. *Behavior Research Methods, Instruments & Computers*, *35*, 136-157.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. RASE, 7(1), 6–10.

Grabe, W. & Stoller, F. L. (2011). Teaching and Researching Reading. Second Edition. Longman: Harlow.

Hamada, M. and Koda, K. (2008). Influence of first language orthographic experience on second language decoding and word leaning. *Language Learning*, *58*, 1-31.

Hamada, M., & Koda, K. (2010). The role of phonological decoding in second language word-meaning inference. *Applied Linguistics*, *31*, 513-531.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. Reading and Writing: An Interdisciplinary Journal, 2, 127–160. https://doi.org/10.1145/504412.504413

Huo, S., & Wang, S. (2017). The effectiveness of phonological-based instruction in English as a foreign language students at primary school level: A research synthesis. *Frontiers in Education*, 1-13.

Hutchison, D. and Styles, B. (2010). *A Guide to Running Randomised Controlled Trials for Educational Researchers*. Slough: NFER. Available from https://www.nfer.ac.uk/media/2114/rct01.pdf.

IPA (International Phonetic Association) (2015) *Full IPA Chart*. Available from http://www.internationalphoneticassociation.org/content/ipa-chart.

Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language learning*, *64*, 160-212.

Koda, K. (2005). Learning to read across writing systems: Transfer, metalinguistic awareness and second-language reading development. In V. Cook & B. Bassetti (Eds.), *Second language writing systems* (pp.311–334). Multilingual Matters.

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. University of Michigan Press.

Leong, C. K., Cheng, P. W., & Tan, L. H. (2005). The role of sensitivity to rhymes, phonemes and tones in reading English and Chinese pseudowords. *Reading and Writing*, *18*, 1-26.

Li, H., & Rao, N. (2000). Parental influences on Chinese literacy development: A comparison of preschoolers in Beijing, Hong Kong and Singapore. *International Journal of Behavioral Development*, *24*, 82–90.

Li. S. (2019). The effects on phonics instruction on L2 phonological decoding and vocabulary learning: an experimental trial on Chinese university EFL learners. Ph.D thesis. University of Oxford.

Li, S., & Woore, R. (2021). The effects of phonics instruction on L2 phonological decoding and vocabulary learning: An experimental study of Chinese EFL learners. *System*, *103*, 102677.

Li, Y. S., & Chen, S. H. (2016). Relative effectiveness of phonological and morphological awareness training on L2 word reading in EFL children. *System*, *60*, 93-104.

Lü, C. (2017). The Roles of Pinyin Skill in English-Chinese Biliteracy Learning: Evidence From Chinese Immersion Learners. *Foreign Language Annals*, *50*, 306-322.

McBride-Chang, C., Lin, D., Liu, P. D., Aram, D., Levin, I., Cho, J. R., Shu, H. & Zhang, Y. (2012). The ABC's of Chinese: Maternal mediation of Pinyin for Chinese children's early literacy skills. *Reading and Writing*, *25*, 283-300.

Nassaji, H. (2014). The role and importance of lower-level processes in second language reading. *Language Teaching*, *47*, 1-37.

National Institute of Child Health and Human Development (NICHD) (2000). Report

of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office. https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf

Odlin, T., (1989). Language Transfer: Cross-linguistic Influence in Language Learning. Cambridge University Press.

Odo, D. M. (2021). A meta-analysis of the effect of phonological awareness and/or phonics instruction on word and pseudo word reading of English as an L2. *SAGE Open*, *11*, 1-13.

Ofsted (Office for Standards in Education, Children's Services and Skills, 2021). Curriculum research review for languages. https://www.gov.uk/government/publications/curriculum-research-review-series-languages.

Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language-learning*. Cambridge University Press.

Rose, J. (2006). Independent review of the teaching of early reading final report. In Department for Education and Skills (Issue March). www.standards.dcsf.gov.uk/phonics/report.pdf

Rowe, K. (2005). Teaching reading: Report and recommendations. National Inquiry

into the Teaching of Literacy. Australian Government, Department of Education, Science and Training. https://research.acer.edu.au/cgi/viewcontent.cgi?article=1004&context=tll_misc

Schiff, R., & Calif, S. (2007). Role of phonological and morphological awareness in L2 oral word reading. *Language Learning*, *57*, 271-298.

Sparks, R., Patton, J., Ganschow, L., & Humbach, N. (2009). Long-term crosslinguistic transfer of skills from L1 to L2. *Language Learning*, *59*, 203-243.

Taft, M., Zhu, X. and Peng, D. (1999). Positional specificity of radicals in Chinese character recognition. *Journal of Memory and Language*, *9*, 182-198.

Woodcock, R. W. (2011). *Woodcock reading Mastery tests: WRMT-III. Circle Pines*. Pearson.

Woore, R. (2009). Beginners' progress in decoding L2 French: Some longitudinal evidence from English modern foreign languages classrooms. *Language Learning Journal*, *37*, 3–18.

Woore, R. (2021). 'Teaching phonics in a second language'. In E. Macaro & R. Woore (eds.), *Debates in Second Language Education* (pp. 222-246). London: Routledge.

Wyse, D., & Bradbury, A. (2022). Reading wars or reading reconciliation? A critical examination of robust research evidence, curriculum policy and teachers' practices for teaching phonics and reading. *Review of Education*, *10*, e3314.

Xu, F., & Ren, P. (2004). The relationship between Chinese children's phonological

awareness and Pinyin skill. *Chinese Journal of Applied Psychology*, *10*, 22–27.

Zhang, S. Z., Georgiou, G. K., Inoue, T., Zhong, W. W., & Shu, H. (2020). Do pinyin and character recognition help each other grow?. *Early Childhood Research Quarterly*, *53*, 476-483.

Zhang, Y. (2016). *National college entrance exam in China: Perspectives on education quality and equity*. Springer.