MACQUARIE
University

Macquarie University PURE Research Management System

This is the author version of an article published as:

Li, L., Marinus, E., Castles, A., Hsieh, M. L., & Wang, H. C. (2021). Semantic and phonological decoding in children's orthographic learning in Chinese. *Scientific Studies of Reading*, 25(4), 319-334.

© 2020. This is an Accepted Manuscript of an article published by Taylor & Francis in *Scientific Studies of Reading* on 04/10/2020, available online https://doi.org/10.1080/10888438.2020.1781863

It is deposited under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Author Accepted Manuscript
DOI:10.1080/10888438.2020.1781863

Semantic and phonological decoding in children's orthographic learning in Chinese

Luan Li[1,2*], Eva Marinus[3], Anne Castles[1,2], Miao-Ling Hsieh[4], Hua-Chen Wang[2,5]

[1] Department of Cognitive Science, Macquarie University

[2] Macquarie University Centre for Reading

[3] Institute for Media and School (IMS), The Schwyz University of Teacher Education

[4] Department of English, National Taiwan Normal University

[5] Department of Educational Studies, Macquarie University

*Corresponding author:

Luan Li

Level 3, Australian Hearing Hub,16 University Avenue, NSW 2113, Australia

luan.li@mq.edu.au

61 452 468 055

Abstract

In this study, we investigated if children build a print-to-meaning connection via the semantic radical – a mechanism we call *semantic decoding* – and its interaction with phonological decoding in orthographic learning of Chinese compound characters. Ninety-two Grade 4 children were taught the pronunciations and meanings of 16 pseudocharacters and were then exposed to their written forms in short stories. Half of the characters contained semantic radicals related to the taught meaning; the other half were unrelated. Half of the children learned the pseudocharacters' regular pronunciations; the other half learned the irregular pronunciations. There was better orthographic learning of regular than irregular pseudocharacters across measures of spelling and orthographic choice. However, the effect of semantic decoding was not as unequivocal. The findings indicate that in Chinese, decoding via phonetic radicals underpins orthographic learning. Whereas, the effect of semantic decoding might depend more on the linguistic features of the radicals and the learning environments.

Key words: semantic decoding, semantic radical, phonological decoding, phonetic radical, orthographic learning, Chinese

Skilled reading relies on well-specified orthographic representations for individual words, and the process of acquiring these representations has been referred to as orthographic learning (Nation & Castles, 2017). One of the most prominent theoretical accounts of orthographic learning is the *self-teaching hypothesis* (Share, 1995), which highlights the role of translating letters to sounds through independent text reading. The importance of the phonological decoding process as a self-teaching mechanism for orthographic learning has now been widely accepted. Even in Chinese, where the orthography is considered not closely related to phonology, phonological decoding via phonetic radicals has been shown to provide direct support for orthographic learning (Li, Wang, Castles, Hsieh, & Marinus, 2018; Li, Li, & Wang, 2019). As written Chinese characters are also connected to meaning via the semantic radical, it is reasonable to hypothesize a *semantic decoding* mechanism comparable to phonological decoding and that children might use it to self-teach. Given the centrality of phonological decoding, it is also possible that semantic decoding plays a complementary role in the formation of novel orthographic representations. In the present study, we present the first controlled investigation of these hypotheses by probing the effect of semantic decoding and its interplay with phonological decoding on Chinese children's orthographic learning.

Orthographic learning in alphabetic languages

The *self-teaching hypothesis* (Share, 1995) postulates that phonological decoding provides a spoken form that binds with the visual form of a written word, which enables children to learn novel words through independent reading. In line with this idea, research has found that correct decoding is associated with success in orthographic learning (Bowey & Miller, 2007; Cunningham, Perry, Stanovich, & Share, 2002; Share, 1999). In addition to supporting forming the orthography-phonology connection, Share (1999) proposed that phonological decoding also draws the reader's attention to the details of the graphemes of the word, thus helping to specify a representation for the word in the orthographic lexicon. This suggestion could explain the findings that orthographic learning was less robust when novel words were surrounded by contextual information that potentially diverts the reader's attention (Landi, Perfetti, Bolger, & Dunlap, 2006) and that a phonological decoding aid presented alongside novel words, albeit providing the correct phonology, did not enhance orthographic learning (Li et al., 2018).

Beyond phonological decoding, another factor associated with orthographic learning is knowledge of word meaning. The role of lexical semantics in word reading is accounted for by the *lexical quality hypothesis* (Perfetti & Hart, 2002), which posits that phonology, semantics and orthography jointly contributes to a word representation in the mental lexicon. Accordingly, there is evidence that providing word semantics during (Ouellette, 2010; Ouellette & Fraser, 2009) or prior to exposure to novel written words (McKague, Pratt, & Johnston, 2001; McKay, Davis, Savage, & Castles, 2008) benefits orthographic learning of novel words in English, although some studies did not find such benefit (e.g., Duff & Hulme, 2012; Landi et al., 2006; Nation, Angell, & Castles, 2007).

While phonological decoding of a word is a sublexical assembly process, semantic knowledge about the word is organized at the lexical level. In alphabetic languages, except for morphologically complex words, there is a lack of systematic print-to-meaning mappings. Hence, the meaning of novel words, particularly monosyllabic words children first learn to read, is often provided "externally" and does not draw attention to its orthography. Therefore, lexical semantics might not contribute to orthographic learning in the way phonological decoding has been shown to. In comparison, written Chinese is more systematically connected to both phonology and semantics at the sublexical level. This raises an interesting potential orthographic learning mechanism of "semantic decoding" in Chinese that is similar to phonological decoding, as the semantic radical in Chinese characters could draw the reader's attention to the form of the characters as well as providing semantic information about them. Below we elaborate on the semantic and phonological decoding system in Chinese.

Semantic and phonological decoding in Chinese

Most Chinese characters are semantic-phonetic compound characters (e.g., 飯 "rice" /fan4/, consisting of a semantic radical on the left (i.e., 飠 "food, to eat") and a phonetic radical on the right (i.e., 反 /fan3/). The semantic radical provides clues to the meaning and the phonetic radical clues to the sound. A compound character can be either transparent or opaque depending on whether its meaning is related to the semantic radical. Likewise, it can be either regular or irregular depending on whether its sound is identical to or different from its phonetic radical (see Table 1 for examples).

Table 1.
Examples of transparent, opaque, regular and irregular compound characters

Semantic radical 飠 "food"		Phonetic radical 反 /fan3/	
Transparent character	Opaque character	Regular character	Irregular character
飯 "rice"	館 "hall, gallery"	販 /fan4/	皈 /gui1/

Although neither the semantic radical nor the phonetic radical is particularly reliable in representing compound characters' meaning and sound (Shu, Chen, Anderson, Wu, & Xuan, 2003), when encountering an unfamiliar character, Chinese children tend to use the semantic radical to deduce its meaning (Chan & Nunes, 1998; Shu & Anderson, 1997) and the phonetic radical to infer its sound (Anderson, Li, Ku, Shu, & Wu, 2003; He, Wang, & Anderson, 2005; Shu, Anderson, & Wu, 2000; Yin & McBride, 2015). For example, Shu and Anderson (1997) selected a set of words that are familiar from children's oral vocabulary but are unfamiliar in written forms (e.g., "pupil of the eye"), and then presented these words in Pinyin, a reliable phonological decoding aid children use to sound out characters. The children were asked to select a written form for the target character from four choices. The choices were compound characters with the same phonetic but different semantic radicals: 瞳, 撞, 僮 and 潼, where the semantic radicals mean "eye", "hand", "people" and "water" respectively. The results showed that children picked 瞳 (whose semantic radical means "eye") far more often than the other choices, suggesting a tendency in using the semantic radical to represent a novel compound character's meaning.

There is also good evidence that children's awareness of the radicals' functions is related to their reading ability (Ho, Ng, & Ng, 2003; Shu, Anderson, & Wu, 2000; Tong, Tong, & McBride, 2017a), as well as their ability to learn to read characters (Packard et al., 2006; Wu et al., 2009). Packard and colleagues (2006) provided children with explicit instruction on decomposing compound characters and using the radicals to represent sound and meaning, and found the instructions improved their character writing ability. This research has established that children can benefit from instructions about the structure of compound characters. Nevertheless, we do not know whether children use the information from radicals to learn novel orthographic forms in a self-teaching environment, without explicit instruction. It is also unclear whether there is a difference in the roles of phonetic and semantic radicals during this learning process. These questions need to be examined by carefully controlled orthographic learning paradigms.

Recent research has used an orthographic learning paradigm to demonstrate that children use the phonetic radical in self-teaching to achieve orthographic learning. Li et al. (2018) taught Grade 2 children the pronunciations and meanings of regular and irregular pseudocharacters (i.e., novel combinations of known semantic and phonetic radicals), and then exposed them to the written forms in stories. The regular characters gained stronger orthographic representations than the irregular characters as measured by spelling and orthographic choice tasks, and this advantage was maintained when reassessed a week later. This study provided strong evidence for a role for phonetic radicals in orthographic learning in a self-teaching environment, yet the contribution of semantic radicals in such condition remains unexplored.

A role for semantic decoding in orthographic learning in Chinese?

There is also evidence that having some semantic knowledge may benefit children's learning to read Chinese words. Zhou et al. (2015) found that teaching 7-and 8-year-olds the meaning and sound of some unknown Chinese words improved their subsequent reading accuracy of the written words relative to teaching the sound only. Nevertheless, learning of the novel words was only measured by reading accuracy in this study, which does not tap the specificity of orthographic representations. Moreover, the role of semantics was explored at the word phrase level (two characters for each word phrase) but not at the sublexical (radical) level.

To explore whether the semantic radical is used to facilitate orthographic learning, Li, Marinus, Yu, Castles and Wang (2019) examined this issue with skilled readers. They asked 41 adult readers of Chinese to read sentences containing 16 pseudocharacters. Half of the characters were transparent, this is, their semantic radicals were related to the meaning conferred by the sentential context (e.g., , of which the semantic radical 日 means "sun", referred to a lighting adjustment system). The other half were opaque, this is, their semantic radicals and the context were unrelated (e.g., , of which the semantic radical ⻊ means "foot", referred to an automatic cooking machine). After five exposures to the pseudocharacters, the participants took written cloze and orthographic decision tasks to assess orthographic learning, and also a definition production task to assess their learning of the print-meaning connections. Li et al., (2019) did not find a transparency effect in the orthographic learning outcomes. Nevertheless, in the definition task, the participants could recall the meanings of the transparent characters better than those of the opaque characters, indicating that semantic decoding might help with learning the orthography-to-semantics mappings.

It is possible that the effect of semantic decoding on orthographic learning may be modulated by the success of phonological decoding. Contemporary computational models of word reading have shown that semantic knowledge mainly contributes to learning to read alphabetic words that allow partial phonological decoding (Plaut, McClelland, Seidenberg, & Patterson, 1996; Pritchard, Coltheart, Marinus, & Castles, 2018). Training studies have also established a semantic benefit on orthographic learning for inconsistent words (McKay et al., 2008; Taylor et al., 2011; Wang et al., 2013). A more direct support for this hypothesis comes from a recent finding that Chinese adults named semantically transparent characters more accurately and more quickly than opaque characters when they had inconsistent pronunciations, and not when they were consistent (Dang, Zhang, Wang, & Yang, 2019). This indicates that the influence of semantic transparency on skilled character recognition may be modulated by phonetic regularity. Given that Li et al. (2019) did not provide phonological information or manipulate phonetic regularity, it is possible that the participants treated all the pseudocharacters as regular and any effect of semantic decoding was thus reduced. It is also unknown whether the results from Li et al. (2019) apply to children, whose incomplete knowledge about radicals might influence their use of them in orthographic learning.

In the present study, we manipulated both semantic transparency and phonetic regularity to investigate the role of semantic decoding and its interplay with phonological decoding in children's orthographic learning. We ask: 1) whether children use semantic decoding to support orthographic learning and forming the print-meaning connections via self-teaching; and 2) whether the effect of semantic decoding is modulated by regularity in the print-sound mappings.

To simulate children's typical reading experience and enable us to manipulate phonological regularity, we taught them the meanings of novel spoken words and then exposed them to the orthography of those words in short stories. We assessed orthographic learning immediately after orthographic exposure and after an approximately 7-day delay. Learning of the orthography-semantics connections was also measured. Based on previous research, we predicted that phonological decoding would support orthographic learning (Li et al., 2018; Li et al., 2019) and semantic decoding would enhance learning the print-meaning mappings (Li et al., 2019). It is possible that semantic decoding would have a stronger effect on the orthographic learning of irregular than regular items. Also, following previous findings (Li et al., 2018), we expected that the overall learning effect would be reduced in a delayed test session 7 days later.

Method

Design

We used a 2 (regularity: regular vs. irregular) × 2 (transparency: transparent vs. opaque) × 2 (testing session: immediate vs. delayed) design. Transparency and testing session were within-group factors. Due to limited number of items we are able to train the children in this learning paradigm, regularity was implemented as a between-group factor. Each class was randomly assigned to either the regular group or the irregular group. This study was preregistered at https://osf.io/ztvy6/, and unless otherwise indicated, we followed the research and analysis protocol as stated in our preregistration.

Participants

This study was approved by the Macquarie University Human Research Ethics Committee. Informed parental consent was obtained for all children. The participants were ninety-two Grade 4 children (49 boys; age *M* = 10 years, *SD* = 4 months), recruited from five classes[1] in three elementary schools in the metropolitan area in Taipei. All were native Chinese speakers without any known language or reading disorders. Their reading ability was assessed with the classroom-based *Chinese Word Recognition Test* (Huang, 2001). Following the standardized protocol in this test, the children were asked to write down Zhuyin for 200 characters arranged in increasing order of difficulty. Like Pinyin, Zhuyin is a transparent phonological coding system that children use to help them read unfamiliar characters in Taiwan. Scoring was discontinued when the child failed to write the correct Zhuyin for 20 consecutive items. The results were normally distributed. The children in this study scored an average of 95 (*SD* = 30), which is higher than the mean score of 83 for fourth graders in Taiwan.

The regular and irregular groups did not differ in age ($t = 1.01$, $p = .314$), but they differed in their standardized reading performance ($t = 8.78$, $p < .001$). The regular group (raw score *M* = 102, *SD* = 27, range = 41 to 165) were on average better readers than the irregular group (raw score *M* = 86, *SD* = 30, range = 26 to 173). This group difference was unexpected but was accounted for in later analyses by including the children's score in the standardized character recognition test as a covariate.

Materials

[1] In order to approximate equal numbers of participants in the regular and irregular groups, we assigned three classes (*N*s = 17, 11, 23) to the regular group and two classes (*N*s = 19, 22) to the irregular group.

Sixteen pseudocharacters were created as targets for the children to learn (twelve experimental items and four fillers) (see Appendix A). They were novel combinations of a semantic radical and a phonetic radical which were of high frequency (over 50 occurrences per million in Ho, 1998). All pseudocharacters had a left-right structure with the semantic radical on the left and the phonetic radical on the right. For example, 飳 comprised a semantic radical 飠 ("food") on the left and a phonetic radical 主 (/zhu3/) on the right. We chose this structure because it is the most frequent character structure in Chinese (Lui, Leung, Law, & Fung, 2010; Shu et al., 2003); therefore, we expected that the participants would use the radicals as cues for meaning and sound.

Each target character was assigned a regular and an irregular pronunciation based on its phonetic radical. The regular pronunciation was the pronunciation of its phonetic radical; the irregular pronunciation was unrelated to the phonetic radical's pronunciation and did not share either the onset or the rime. Take the target 飳 for example, its regular pronunciation was /zhu3/ - the same as its phonetic radical 主 /zhu3/; the irregular pronunciation /tao2/ was unrelated.

Following Li et al. (2018), the target characters were embedded in three-character words (超x機 "super X machine", with the target pseudocharacter inserted in the middle), which referred to novel inventions. Eight of the inventions were adapted from Wang, Castles, Nickels, and Nation (2011), with another eight from Mimeau, Ricketts, and Deacon (2017). Because all the words started and ended with the same characters, the assumption was that the children would focus on and learn the middle pseudocharacters during the training.

Each word was assigned a transparent and an opaque meaning based on its semantic radical. For example, in the transparent condition, the middle pseudocharacter in 超飳機 has a semantic radical 飠, which means "food" or "to eat", and word referred to "a food processor that can take out the food people do not like". In the opaque condition, it referred to "a machine that allows people to walk on the wall."

For each semantic radical, we calculated a consistency value to index the percentage of existing characters with the semantic radical that have related/transparent meanings. The characters included for calculation are amongst the most frequent 3,091 characters in Ho (1998). For example, among the twelve characters with the semantic radical 飠 ("food"), seven are transparent, having a "food"-related meaning[2]: 飽 ("full of food"), 餅 ("pancake"), 飯 ("rice"), 餓 ("hungry"), 飢 ("starving"), 餵 ("feed") and 飲 ("drink"). Therefore, 飠 is .58 (7/12) consistent in being associated with the transparent meanings in existing characters. The semantic radicals we selected to create the pseudocharacters had a mean consistency value of .65, ranging from .5 to .95.

We also conducted a radical knowledge test at the end of the experiment to ensure that the children were indeed familiar with the radicals we selected. The children were asked to name the phonetic radicals by writing down their pronunciations with Zhuyin. They were also asked to choose one from four pictures that might represent the meaning or the semantic category of the semantic radical – a test adapted from Ho, Ng and Ng (2003). The correct choice pictured a real object that was similar to the inventions used in this study in both appearance and function. The results showed that the children had sound knowledge of the phonetic radicals' pronunciations (accuracy M = .99, SD = .02) and the semantic radicals' meanings (accuracy M = .85, SD = .20).

All participants learned all sixteen items. The regular group (N= 51) learned the regular pronunciations for the twelve experimental items and the irregular pronunciations for the four fillers, and vice versa for the irregular group (N= 41). Fillers were included so that

[2] The first author and a research assistant independently judged the transparency of the characters used for calculation. Disagreements were resolved by another native Mandarin Chinese speaker.

the children would not strategically read all targets as either regular or irregular. Also, we wanted to approximate the children's real-world reading experience, where regular and irregular characters are both common. However, because the fillers were not matched on the orthographic and linguistic features in the same way as the targets, they were not included in the analyses. The sixteen items were grouped into two lists, and matched on visual complexity (i.e., the total number of strokes), $t(14) = .45$, $p = .662$, and semantic radical consistency, $t(14) = 1.26$, $p = .228$. Half of the participants in each group learned the transparent meanings for one list and the opaque meanings for the other list, and this was counterbalanced across groups such that all items appeared in all four conditions.

Procedure

Training and testing took place over three weeks in the classrooms of the participating schools. All training and testing were group/class-based and were carried out by the first author and two research assistants following a script. The procedure was similar to Wang and colleagues (2012; 2011) and Li et al. (2018), which included oral vocabulary learning prior to print exposure. We adopted this paradigm because it allows us to manipulate phonetic regularity in a self-teaching environment where no instruction is required during orthographic learning. Also, it is a more naturalistic reflection of children's literacy development, given that children are often familiar with a word in their oral vocabulary before seeing it in print. We describe the details of the procedure below.

Oral vocabulary learning (3 days in Week 1). The children were taught vocabulary of the target words via pictures and oral explanations, including functions and key features of the inventions presented in the pictures (see Appendix B for an example). Half of the items from each of the conditions were taught on the first day and the other half on the second day.

On the third day, all sixteen items were trained together. Training took approximately 20 minutes per day.

Orthographic exposure (2 days in Week 2). On each orthographic exposure day, before the children saw the target characters' written forms, we first administered a picture naming test to measure oral vocabulary learning.

Picture naming. The children were given an answer sheet with pictures of the inventions and asked to name the pictures by writing down their pronunciations in Zhuyin.

Orthographic learning. Next, they were each given a story book and were asked to silently read the stories with the trained words in them (see Appendix B for an example). No feedback was provided. Before they started reading, they were instructed to use what they had learned earlier about the invention names to help them with the reading. After they finished reading, we administered two comprehension questions to ensure that they had focused on reading. The children were exposed to eight stories on the first day and to another eight on the second day.

Immediate test (2 days in Week 2). On each orthographic exposure day, after the children finished reading, the story books were put away. Then spelling and orthographic choice tasks were conducted to measure orthographic learning. Word-picture matching was also administered to assess learning of the print-meaning connections. The children wrote down their responses on answer sheets, which were given to them before each task and were collected after each task. Unlimited time was given to produce the responses.

Spelling. The experimenter read out the target words, and the children were asked to try to write down the missing target characters exactly as they had seen in the stories. The correct response was the correct combination of the semantic and phonetic radicals in the correct position for a target character.

Orthographic choice. The children were asked to choose the word they had seen in the stories from five choices presented on the answer sheets (see Appendix A for a full list). The choices included the target, a phonological foil, a semantic foil and two visual distractors, which were also all pseudocharacters. All foils were either pseudocharacters or low-frequency characters unlikely known by the children. The phonological foil had the same semantic radical as the target. Its phonetic radical was either a homophone of the target's pronunciation or differed from it only in tone. The semantic foil had the same phonetic radical but a different semantic radical as the target. The meaning of its semantic radical was related to the invention the target referred to. Take  for an example, in the transparent condition, the foil's semantic radical was 口, which means "mouth" and was related to the food processer. In the opaque condition, where the target referred to a wall-walking machine, the foil's semantic radical was ⻊ , which means "foot". One visual distractor had a different but similar-looking semantic radical, and the other had a different but similar-looking phonetic radical to the target character.

Word-picture matching. Finally, the children were given answer sheets with all the target written words and the invention pictures. The task was to match the words with the corresponding invention pictures. The correct response was the correct matching for a target word.

Delayed test (1 day in week 3). Approximately seven days after initial orthographic exposure, the children took spelling, orthographic choice and word-picture matching tasks again in the same order, to assess how well the learning had been retained.

Results

Logistic mixed effects regression models with maximal structure (Barr, Levy, Scheepers, & Tily, 2013) were fitted using the lme4 package (Bates, Mächler, Bolker, &

Walker, 2014) in the R computing environment (R Core Team, 2013). Planned contrasts were computed using the lsmeans package (Length, 2016). Separate models were run for each dependent variable: accuracies in picture naming, spelling, orthographic choice and word-picture matching. Fixed effects included transparency, regularity, testing session and their interactions. As picture naming was only administered in the immediate sessions, testing session was not implemented as a fixed effect for this model. Raw score from the standardized reading test was initially included as a covariate in all models, but was removed from the final models because it reduced model fits, and did not produce any significant effect on the outcome measures. By-subject and by-item random intercepts and slopes were included for each fixed main effect and interaction. When a model failed to converge, following Barr et al.'s (2013) suggestion, we dropped the random slopes associated with smaller variance explained until the model converged (see Appendix C for the final model).

In the following section, we first report overall outcomes of oral vocabulary learning and then learning of the orthographic form and meaning of the target items. Next, we report the effects of transparency and regularity on orthographic learning, followed by their effects on learning the print-meaning connections. Descriptive statistics are presented in Table 2.

How well did the children learn the oral vocabulary?

Picture naming. The children's performance in this task suggests that they had acquired substantial oral vocabulary of the novel words. There was no main effect of transparency, $\beta = .22$, $SE = .22$, $z = 1.01$, $p = .313$) or interaction between transparency and regularity, $\beta = -.41$, $SE = .33$, $z = -1.24$, $p = .216$. However, the main effect of regularity was significant, $\beta = 1.59$, $SE = .36$, $z = 4.46$, $p < .001$). The regular group named the pictures more accurately than the irregular group.

Table 2.

Mean proportions of correct responses in the learning tasks (standard deviations in parentheses).

		Transparency		
Task	Regularity	Transparent	Opaque	Average
Picture naming	Regular	.85 (.36)	.87 (.34)	.86 (.35)
	Irregular	.69 (.46)	.65 (.48)	.67 (.47)
	Average	.78 (.42)	.77 (.42)	.77 (.42)
(Immediate test)				
Spelling	Regular	.66 (.47)	.67 (.47)	.67 (.47)
	Irregular	.26 (.44)	.18 (.38)	.22 (.42)
	Average	.49 (.50)	.45 (.50)	.47 (.50)
Orthographic choice	Regular	.91 (.28)	.89 (.32)	.90 (.30)
	Irregular	.67 (.47)	.65 (.48)	.66 (.47)
	Average	.81 (.40)	.78 (.41)	.79 (.40)
Word-picture matching	Regular	.90 (.30)	.90 (.30)	.90 (.30)
	Irregular	.39 (.49)	.35 (.48)	.37 (.48)
	Average	.67 (.47)	.65 (.48)	.66 (.47)
(Delayed test)				
Spelling	Regular	.55 (.50)	.47 (.50)	.51 (.50)
	Irregular	.09 (.29)	.07 (.25)	.08 (.27)
	Average	.34 (.48)	.29 (.45)	.31 (.46)
Orthographic choice	Regular	.86 (.34)	.80 (.40)	.83 (.37)
	Irregular	.59 (.49)	.52 (.50)	.55 (.50)
	Average	.74 (.44)	.67 (.47)	.71 (.46)
Word-picture matching	Regular	.83 (.37)	.84 (.36)	.84 (.37)
	Irregular	.18 (.39)	.08 (.27)	.13 (.34)
	Average	.54 (.50)	.49 (.50)	.52 (.50)

This between-group difference was unexpected as the classes were randomly assigned to the two groups. It seems that the regular group were not only on average better readers but were also better at oral vocabulary learning. This is not surprising given that there is a known reciprocal relationship between Chinese children's reading ability and oral language skills (e.g., Hulme, Zhou, Tong, Lervåg, & Burgoyne, 2018). Given that this between-group difference might confound any regularity effect on the orthographic learning outcomes, we accounted for this by conducting additional analyses with picture naming accuracy included in the models as a fixed covariate. The results were similar and are reported below in the last section of Results. It is also important to note that the main research question under investigation was the effect of semantic transparency on orthographic learning, which is manipulated as a within-subject factor and should not be affected by this between-group difference.

How well did the children learn the pseudocharacters?

The children's overall performance was comparable to previous research utilising the same training procedure (Li et al., 2018; Wang et al., 2012, 2013). Averaged across testing sessions, the participants spelled 38.8% of the items correctly. Of their errors, 71.58% were non-responses and 28.42% were spelling errors, most of which were the spelling of a character homophonic to the target. In orthographic choice, the children were on average 74.68% accurate across sessions. Of their erroneous choices, 24.46%, 31.83%, 32.91% and 10.79% were phonological, semantic, visual and non-responses respectively. The main effect of testing session was significant across the measures (spelling: $\beta = 1.40$, $SE = .33$, $z = 4.29$, $p < .001$; orthographic choice: $\beta = .77$, $SE = .28$, $z = 2.72$, $p = .007$; word-picture matching: $\beta = 2.18$, $SE = .41$, $z = 5.28$, $p < .001$). The results suggest that the children had acquired some orthographic and semantic knowledge of the target characters and that this knowledge was partly retained over time.

Does semantic decoding influence orthographic learning and does it interact with phonological decoding?

Spelling. The effect of transparency on spelling was not significant ($\beta = .47$, *SE* = .34, $z = 1.40$, $p = .161$), although numerically, spelling accuracy was higher for transparent than for opaque items. The interaction between transparency and regularity was marginally significant, $\beta = -.66$, *SE* = .34, $z = -1.97$, $p = .049$. However, there was also a significant three-way interaction between transparency, regularity and testing session, $\beta = 1.09$, *SE* = .53, $z = 2.04$, $p = .042$. Further analysis revealed that the transparency by regularity interaction was not significant in either the immediate, $\beta = -.16$, *SE* = .08, $z = -1.97$, $p = .098$, or the delayed session, $\beta = .11$, *SE* = .11, $z = 1.01$, $p = .310$. Hence, the three-way interaction was caused by the transparency by regularity interaction going in opposite directions in the two testing sessions, as shown by the coefficients. There was a significant main effect of regularity, $\beta = 3.60$, *SE* = .48, $z = 7.47$, $p < .001$. Spelling accuracy was higher for regular than irregular items.

Orthographic choice. Despite a numeric difference between transparent and opaque items, the main effect of transparency was not significant, $\beta = .22$, *SE* = .28, $z = .78$, $p = .435$,. The interaction between transparency and regularity was not significant either, $\beta = .16$, *SE* = .44, $z = .35$, $p = .726$. The effect of regularity was significant, $\beta = 2.03$, *SE* = .49, $z = 4.13$, $p < .001$. The accuracy for orthographic choice was significantly higher for regular than for irregular items. None of the interactions involving testing session was significant (ps > .05).

Does semantic decoding influence learning the orthography-semantics connections and does it interact with phonological decoding?

Word-picture matching. The results demonstrated a similar pattern as for the spelling and orthographic choice outcomes. There was no main effect of transparency, $\beta = .20$, $SE = .22$, $z = .91$, $p = .363$, nor an interaction between transparency and regularity, $\beta = -.25$, $SE = .36$, $z = -.70$, $p = .487$. The three-way interaction between transparency, regularity and testing session was marginally significant, $\beta = -1.00$, $SE = .53$, $z = -1.88$, $p = .060$. The interaction between transparency and testing session was also significant, $\beta = .84$, $SE = .38$, $z = 2.19$, $p = .029$. Further analysis revealed a significant transparency effect in the irregular group in the delayed session, $\beta = 1.00$, $SE = .31$, $z = 3.25$, $p = .005$. However, this result needs to be interpreted with caution due to a floor effect. The main effect of regularity was also significant, $\beta = 3.52$, $SE = .39$, $z = 9.01$, $p < .001$. The accuracy was significantly higher for the regular than for the irregular items.

Does oral vocabulary influence learning of the orthography and meaning?

This question was not preregistered. We included this post-hoc analysis due to the unexpected group difference in oral vocabulary learning. To explore this question, we conducted an additional set of analyses with the immediate test outcomes and included picture naming accuracy as a fixed covariate. The results demonstrated the same pattern as the previous analyses in all tasks. Additionally, there was a significant main effect of picture naming accuracy (spelling: $\beta = .84$, $SE = .24$, $z = 3.44$, $p < .001$; orthographic choice, $\beta = .84$, $SE = .24$, $z = 3.59$, $p < .001$; word-picture matching: $\beta = .97$, $SE = .22$, $z = 4.38$, $p < .001$).

Discussion

In this study with Grade 4 Chinese-speaking children, we tested the hypothesis that semantic decoding supports the formation of orthographic representations of novel Chinese compound characters and that its contribution is modulated by phonological decoding.

Specifically, following previous research, we predicted better orthographic learning for regular than for irregular characters (Li et al., 2018), and we hypothesized that the transparency effect might be stronger for irregular characters than for regular characters. We taught the children oral vocabulary and then asked them to read short stories containing pseudocharacters. Orthographic learning was measured with spelling and orthographic choice immediately after reading and after a 7-day delay. We also measured learning of the meaning of the target items to see whether semantic decoding would assist with learning the orthography-semantics mappings.

After 4 training sessions, the children were able to associate the sound and the meaning of the target words more than 75% of the time, which suggests that they had acquired some oral vocabulary for the words. The results from the measures of orthographic learning – spelling and orthographic choice – showed evidence of robust orthographic learning both immediately after exposure and after a 7-day delay. Word-picture matching outcomes also showed evidence of acquisition of the written words' meaning. The children's overall performance was comparable to previous orthographic learning research that also provided oral vocabulary training (Li et al., 2018; Wang et al., 2012, 2011).

We found a significant effect of phonetic regularity in all tasks for both testing sessions. Regular characters were spelled, identified and named more accurately than irregular characters, which is consistent with a previous demonstration that children use the phonetic radical to phonologically decode and to learn to read novel characters in self-teaching (Li et al., 2018; Li et al., 2019). Interestingly, word-picture matching accuracy was also higher when the characters had regular pronunciations, suggesting that phonological decoding might also benefit learning the characters' meaning. This is not surprising, given that the fluent retrieval of word meanings is contingent on well-specified orthographic representations (Perfetti & Hart, 2002), which is supported by phonological decoding.

Previous orthographic learning research has rarely examined the semantic aspect of word learning. Although several studies have shown that children can also acquire knowledge of the print-meaning mappings during self-teaching (Deacon, Mimeau, Chung, & Chen, 2019; Ricketts, Bishop, Pimperton, & Nation, 2011), they did not address whether phonological decoding could facilitate learning the written words' meaning. Here, we found some evidence to suggest that this might be the case. This finding, however, needs to be replicated in future confirmatory studies.

In contrast to phonological decoding, semantic decoding does not seem to be used as a self-teaching mechanism to support orthographic learning. Semantic transparency did not significantly affect spelling or orthographic choice, although there was a numerical trend for an effect. This is consistent with findings of Li et al. (2019) with adult readers. They hypothesized that this might be because the sentential context provided a top-down support to derive the meaning of the character, reducing the necessity to rely on the semantic radical to assist processing the character in text. In contrast, studies that have shown a transparency effect on skilled character reading (Chen & Weekes, 2004; Li & Chen, 1999; Wang, Pei, Wu, & Su, 2017; Williams & Bever, 2010) and on learning the meaning of novel characters in learners of Chinese as a foreign language (Nguyen, Zhang, Li, Wu, & Cheng, 2017; Zhang, Li, Dong, Xu, & Sholar, 2015) typically presented characters in isolation. Another recent study (Li et al., 2019) that demonstrated a transparency effect on orthographic learning (with a small effect size $d = .172$ in orthographic choice) used a paradigm that encouraged attention to the semantic radicals by providing them as stand-alone characters at the beginning of the learning session. In these cases, readers are more likely to focus on the orthography, in particular the radicals, and therefore make use of them.

Semantic decoding also did not function as a compensatory mechanism to phonological decoding in supporting the children's orthographic learning. For the irregular characters (e.g., 飳 /tao2/), even though the phonetic radical (e.g., 主 /zhu3/) provided unreliable cues to be used for phonological decoding, the children did not turn to the semantic radical to learn their orthographic representations. As mentioned above, this might also be due to the presence of the story context, reducing the need to use the semantic radical. It should be noted that this result is not fully comparable to findings in English that semantics assists orthographic learning, particularly for irregular words (McKay et al., 2008; Taylor et al., 2011; Wang et al., 2013). As we have stated earlier, semantic decoding in Chinese is a sublexical process more comparable to phonological decoding. The current findings thus do not address whether lexical semantic knowledge benefits orthographic learning in Chinese as it does in English.

Another possible explanation is that whether a semantic radical would be used for learning novel characters depends on some linguistic features of the radical such as consistency, imageability and neighborhood density. For instance, research has shown that semantic consistency influences skilled character reading (Chen & Weekes, 2004; Feldman & Siok, 1999). In the present study, the semantic radicals we used to create the pseudocharacters ranged in consistency (0.5 ~ 0.95). The rationale was to simulate children's typical learning experience, as most semantic radicals they encounter are not highly consistent. In addition, in the opaque condition, highly-consistent semantic radicals would appear very unnatural to them. Future studies should explore this possibility and examine which features of semantic radicals may affect children's use of them in learning to read.

We also found that the semantic radical did not seem to contribute to learning the print-meaning connections, which seems at odds with what was found by Li et al. (2019). However, a noteworthy difference between Li et al. (2019) and our study is that we provided oral vocabulary training before print exposure, and they did not. Therefore, it might be that in the present study, when the children saw the pseudocharacters in print, they used their extant oral vocabulary acquired from the pre-exposure phase to build the print-meaning links and therefore did not need to rely on the semantic radical. Our post-hoc analyses showing that oral vocabulary facilitated learning the items' meaning has provided some support for this suggestion. This would also align with Tong, Tong and McBride (2017)'s finding that the relationship between children's knowledge about the semantic radical and reading ability is mediated by their vocabulary knowledge. They proposed that the semantic radical facilitates the acquisition of semantic knowledge for unfamiliar written characters, which in turn contributes to reading it. This prediction should be tested in future research.

The centrality of phonological decoding in orthographic learning is largely rationalized by the observation that the alphabetic writing systems represent phonology more systematically than semantics. Yet we have shown that even in Chinese, which features a relatively equal degree of systematicity in print-meaning and -sound connections, phonological decoding is still the primary underpinning of orthographic learning. This finding fits broadly with the universal phonological principle (Perfetti, Zhang, & Berent, 1992), which claims that phonological processes are pervasive in reading across writing systems. On the other hand, our findings suggest that the effect of semantic radicals in orthographic learning of Chinese characters is not unequivocal. It might depend on the linguistic features of the radicals and the learning environments.

Acknowledgements

This research is partly supported by grants from Australian Research Council (grant number DP150100419) to AC. We are grateful to the staff and students at Taipei Municipal Yucheng Elementary School, Taipei Municipal Zhongyi Elementary School and Taipei Municipal Hu Shih Elementary School for participating in this research. We also thank Peiyu Hsieh, Joyce Jang and Wei-Peng Lin for their assistance with data collection. The authors declare no conflict of interest.

References

Anderson, R. C., Li, W., Ku, Y.-M., Shu, H., & Wu, N. (2003). Use of partial information in learning to read Chinese characters. *Journal of Educational Psychology*, *95*(1), 52–57.

Barr, D., Levy, R., Scheepers, C., & Tily, H. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *ArXiv Preprint ArXiv:1406.5823*.

Bowey, J., & Miller, R. (2007). Correlates of orthographic learning in third-grade children's silent reading. *Journal of Research in Reading*, *30*(2), 115–128.

Chan, L., & Nunes, T. (1998). Children's understanding of the formal and functional characteristics of written Chinese. *Applied Psycholinguistics*, *19*(01), 115–131.

Chen, M., & Weekes, B. (2004). Effects of semantic radicals on Chinese character categorization and character decision. *Chinese Journal of Psychology*, *46*(2), 181–196.

Cunningham, A. E., Perry, K. E., Stanovich, K. E., & Share, D. L. (2002). Orthographic learning during reading: Examining the role of self-teaching. *Journal of Experimental Child Psychology*, *82*(3), 185–199.

Dang, M., Zhang, R., Wang, X., & Yang, J. (2019). The interaction between phonological and semantic processing in reading Chinese characters. *Frontiers in Psychology*, *9*, 2748.

Deacon, S. H., Mimeau, C., Chung, S. C., & Chen, X. (2019). Young readers' skill in learning spellings and meanings of words during independent reading. *Journal of Experimental Child Psychology*, *181*, 56–74.

Duff, F. J., & Hulme, C. (2012). The role of children's phonological and semantic knowledge in learning to read words. *Scientific Studies of Reading*, *16*(6), 504–525.

Feldman, L., & Siok, W. (1999). Semantic radicals contribute to the visual identification of Chinese characters. *Journal of Memory and Language*, *40*(4), 559–576.

He, Y., Wang, Q., & Anderson, R. C. (2005). Chinese children's use of subcharacter information about pronunciation. *Journal of Educational Psychology*, *97*(4), 572–579.

Ho, C. S.-H., Ng, T.-T., & Ng, W.-K. (2003). A "radical" approach to reading development in Chinese: The role of semantic radicals and phonetic radicals. *Journal of Literacy Research*, *35*(3), 849–878.

Ho, H. H. (1998). 现代汉语常用字频率统计 [Chinese character Frequency, a trans-regional diachronic survey]. Retrieved from http://humanum.arts.cuhk.edu.hk/Lexis/chifreq/

Huang, H.-S. (2001). *中文年級認字量表[Graded Chinese Character Recognition Test]*. Taipei: Psychological Publishing Company.

Hulme, C., Zhou, L., Tong, X., Lervåg, A., & Burgoyne, K. (2019). Learning to read in Chinese: Evidence for reciprocal relationships between word reading and oral language skills. *Developmental Science*, *22*(1), e12745.

Landi, N., Perfetti, C., Bolger, D., & Dunlap, S. (2006). The role of discourse context in developing word form representations: A paradoxical relation between reading and learning. *Journal of Experimental Child Psychology*, *94*(2), 114–133.

Length, R. V. (2016). Least-squares means: the R package lsmeans. *Journal of Statistical Software*, *69*(1), 1–33.

Li, H., & Chen, H.-C. (1999). Radical processing in Chinese character recognition: Evidence from lexical decision. *Psychologia*, *42*(4), 199–208.

Li, L., Marinus, E., Castles, A., Yu, L., & Wang, H.-C. (2019). Eye-tracking the effect of semantic decoding on orthographic learning in Chinese. Retrieved from psyarxiv.com/ekxd6

Li, L., Wang, H.-C., Castles, A., Hsieh, M.-L., & Marinus, E. (2018). Phonetic radicals, not phonological coding systems, support orthographic learning via self-teaching in Chinese. *Cognition*, *176*, 184–194.

Li, Y., Li, H., & Wang, M. (2019). The roles of phonological recoding, semantic radicals and writing practice in lrthographic Learning in Chinese. *Scientific Studies of Reading*, 1–12.

Lui, H.-M., Leung, M.-T., Law, S.-P., & Fung, R. S.-Y. (2010). A database for investigating the logographeme as a basic unit of writing Chinese. *International Journal of Speech-Language Pathology*, *12*(1), 8–18.

McKague, M., Pratt, C., & Johnston, M. (2001). The effect of oral vocabulary on reading visually novel words: A comparison of the dual-route-cascaded and triangle frameworks. *Cognition*, *80*(3), 231–262.

McKay, A., Davis, C., Savage, G., & Castles, A. (2008). Semantic involvement in reading aloud: evidence from a nonword training study. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *34*(6), 1495–1517.

Mimeau, C., Ricketts, J., & Deacon, S. H. (2017). The role of orthographic and semantic learning in word reading and reading comprehension. *Scientific Studies of Reading*, *22*(5), 384–400.

Nation, K., Angell, P., & Castles, A. (2007). Orthographic learning via self-teaching in children learning to read English: Effects of exposure, durability, and context. *Journal of Experimental Child Psychology*, *96*(1), 71–84.

Nation, K., & Castles, A. (2017). Putting the learning into orthographic learning. In K. Cain, D. Compton, & R. Parrila (Eds.), *Theories of reading development* (pp. 147–167). John Benjamins Publishing Company.

Nguyen, T. P., Zhang, J., Li, H., Wu, X., & Cheng, Y. (2017). Teaching Semantic Radicals Facilitates Inferring New Character Meaning in Sentence Reading for Nonnative Chinese Speakers. *Frontiers in Psychology*, *37*(8), 1846.

Ouellette, G. (2010). Orthographic learning in learning to spell: The roles of semantics and type of practice. *Journal of Experimental Child Psychology*, *107*(1), 50–58.

Ouellette, G., & Fraser, J. R. (2009). What exactly is a yait anyway: The role of semantics in orthographic learning. *Journal of Experimental Child Psychology*, *104*(2), 239–251.

Packard, J. L., Chen, X., Li, W., Wu, X., Gaffney, J. S., Li, H., & Anderson, R. C. (2006). Explicit instruction in orthographic structure and word morphology helps Chinese children learn to write characters. *Reading and Writing*, *19*(5), 457–487.

Perfetti, C. A., Zhang, S., & Berent, I. (1992). Reading in English and Chinese: Evidence for a “universal” phonological principle. *Advances in Psychology*, *94*, 227–248.

Perfetti, C., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 67–86). John Benjamins Publishing Company.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychological Review*, *103*(1), 56–115.

Pritchard, S. C., Coltheart, M., Marinus, E., & Castles, A. (2018). A computational model of the self-teaching hypothesis based on the dual-route cascaded model of reading. *Cognitive Science*, *42*, 722–770.

Ricketts, J., Bishop, D. V. M., Pimperton, H., & Nation, K. (2011). The role of self-teaching in learning orthographic and semantic aspects of new words. *Scientific Studies of Reading*, *15*(1), 47–70.

Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition*, *55*(2), 151–218.

Share, David L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, *72*(2), 95–129.

Shu, H., & Anderson, R. (1997). Role of radical awareness in the character and word acquisition of Chinese children. *Reading Research Quarterly*, *32*(1), 78–89.

Shu, H., Anderson, R. C., & Wu, N. (2000). Phonetic awareness: Knowledge of orthography-phonology relationships in the character acquisition of Chinese children. *Journal of Educational Psychology*, *92*(1), 56–62.

Shu, H., Chen, X., Anderson, R. C., Wu, N., & Xuan, Y. (2003). Properties of school Chinese: Implications for learning to read. *Child Development*, *74*(1), 27–47.

Team, R. C. (2013). R: A language and environment for statistical computing. Retrieved from http://www.r-project.org/

Tong, X.-H., Tong, X.-L., & McBride, C. (2017). Radical sensitivity is the key to understanding Chinese character acquisition in children. *Reading and Writing*, *30*(6), 1251–1265.

Tong, X., Tong, X., & McBride, C. (2017). Unpacking the relation between morphological awareness and Chinese word reading: Levels of morphological awareness and vocabulary. *Contemporary Educational Psychology*, *48*, 167–178.

Wang, H.-C., Castles, A., & Nickels, L. (2012). Word regularity affects orthographic learning. *The Quarterly Journal of Experimental Psychology*, *65*(5), 856–864.

Wang, H.-C., Castles, A., Nickels, L., & Nation, K. (2011). Context effects on orthographic learning of regular and irregular words. *Journal of Experimental Child Psychology*, *109*(1), 39–57.

Wang, X., Pei, M., Wu, Y., & Su, Y. (2017). Semantic radicals contribute more than phonetic radicals to the recognition of Chinese phonograms: Behavioral and ERP evidence in a factorial study. *Frontiers in Psychology*, *8*, 2230.

Williams, C., & Bever, T. (2010). Chinese character decoding: a semantic bias? *Reading and Writing*, *23*(5), 589–605.

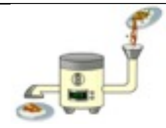
Wu, X., Anderson, R. C., Li, W., Wu, X., Li, H., Zhang, J., … Jiang, W. (2009). Morphological awareness and Chinese children’s literacy development: An intervention study. *Scientific Studies of Reading*, *13*(1), 26–52.

Yin, L., & McBride, C. (2015). Chinese kindergartners learn to read characters analytically. *Psychological Science*, *26*(4), 424–432.

Zhang, J., Li, H., Dong, Q., Xu, J., & Sholar, E. (2015). Implicit use of radicals in learning characters for nonnative learners of Chinese. *Applied Psycholinguistics*, *37*(8), 1–21.

Zhou, L.-L., Duff, F. J., & Hulme, C. (2015). Phonological and semantic knowledge are causal influences on learning to read words in Chinese. *Scientific Studies of Reading*, *19*(6), 409–418.

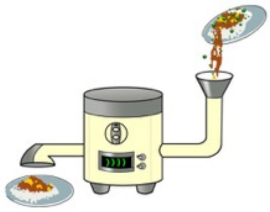
Appendix A. Target pseudocharacters and items in the orthographic choice task. (SR: semantic radical; PR: phonetic radical)

	Target pseudocharacter	Semantic foil		Phonological foil		Visual distractor		Invention picture (in transparency condition)
		Transparent	Opaque	Regular	Irregular			
List A	⿰食主 (SR: food; PR: \zhu3\)	⿰口主 (SR: mouth)	⿰足主 (SR: foot)	⿰食朱 (PR: \zhu1\)	⿰食匋 (PR: \tao2\)	⿰食壬	⿰馬主	
	⿰扌文 (SR: hand; PR: \wen2\)	⿰片文 (SR: card)	⿰冫文 (SR: ice)	⿰扌聞 (PR: \wen2\)	⿰扌孝 (PR: \xiao4\)	⿰扌叉	⿰木文	
	⿰言由 (SR: speak; PR: \you2\)	⿰犭由 (SR: animal)	⿰火由 (SR: fire)	⿰言尤 (PR: \you2\)	⿰言夫 (PR: \fu1\)	⿰言田	⿰足由	
	⿰土台 (SR: dirt; PR: \tai2\)	⿰木台 (SR: wood)	⿰日台 (SR: sun)	⿰土太 (PR: \tai4\)	⿰土克 (PR: \ke4\)	⿰土呂	⿰王台	
	⿰衤平 (SR: cloth; PR: \ping2\)	⿰糸平 (SR: silk)	⿰氵平 (SR: water)	⿰衤并 (PR: \ping2\)	⿰衤暴 (PR: \bao4\)	⿰衤米	⿰木平	
	⿰目皮 (SR: eye; PR: \pi2\)	⿰亻皮 (SR: people)	⿰金皮 (SR: gold)	⿰目匹 (PR: \pi2\)	⿰目各 (PR: \ge4\)	⿰目反	⿰口皮	
List B	⿰足里 (SR: foot; PR: \li3\)	⿰革里 (SR: leather)	⿰食里 (SR: food)	⿰足立 (PR: \li4\)	⿰足炎 (PR: \yan2\)	⿰足果	⿰石里	
	⿰冫半 (SR: ice; PR: \ban4\)	⿰氵半 (SR: water)	⿰言半 (SR: speak)	⿰氵般 (PR: \ban1\)	⿰冫夆 (PR: \feng1\)	⿰冫木	⿰王半	

	烄 (SR: fire; PR: \jiao1\)	[illegible] (SR: wind)	[illegible] (SR: hand)	[illegible] (PR: \jiao3\)	[illegible] (PR: \qi2\)	[illegible]	[illegible]	
	泩 (SR: water; PR: \sheng1\)	[illegible] (SR: ice)	[illegible] (SR: dirt)	汫 (PR: \sheng1\)	湄 (PR: \mei2\)	[illegible]	[illegible]	
	[illegible] (SR: gold; PR: \yang1\)	[illegible] (SR: sun)	眻 (SR: eye)	鈌 (PR: \yang1\)	鉭 (PR: \dan4\)	[illegible]	[illegible]	
	[illegible] (SR: sun; PR: \tong1\)	烔 (SR: light)	[illegible] (SR: cloth)	曈 (PR: \tong2\)	晥 (PR: \wan2\)	[illegible]	[illegible]	
Fillers	[illegible] (SR: wheat; PR: \he2\)	[illegible] (SR: knife)	[illegible] (SR: body)	[illegible] (PR: \he2\)	[illegible] (PR: \jia1\)	[illegible]	[illegible]	
	[illegible] (SR: body; PR: \qing1\)	腈 (SR: moon)	[illegible] (SR: wheat)	[illegible] (PR: \qing1\)	[illegible] (PR: \cheng2\)	[illegible]	[illegible]	
	[illegible] (SR: music; PR: \xi1\)	[illegible] (SR: bird)	[illegible] (SR: fish)	[illegible] (PR: \xi1\)	[illegible] (PR: \shi4\)	[illegible]	[illegible]	
	魬 (SR: fish; PR: \fan3\)	汳 (SR: water)	[illegible] (SR: music)	[illegible] (PR: \fan2\)	[illegible] (PR: \lan2\)	[illegible]	[illegible]	

Appendix B. An example of the pictures and procedure in the oral vocabulary learning phase, and the stories that the children read in the orthographic exposure phase.

<u>Oral vocabulary learning phase</u>



- The experimenter showed the children the picture and taught them the name of the invention in spoken form (e.g., This is a 超飳機 /chao1 zhu3 ji1/). The children repeated the name.
- The experimenter described the semantic features of the invention (e.g., 超飳機 /chao1 zhu3 ji1/) is used to take out the food you don't like from a meal. It has a tube and two open ends). The children were asked to repeat the invention's name again.
- The children repeated the invention's name and its function.
- Picture-naming task. The children were asked to recall the name of the invention when presented with a picture (in random order). Feedback was provided.
- In Sessions 3 and 4, the children were also asked to complete twelve sentences like "If I had a 超飳機 /chao1 zhu3 ji1/, I would use it to …" or "If I want to take out the food I don't like from a meal, then I should use …"

<u>Orthographic exposure phase</u>

小明的午餐里有豌豆，可是小明不喜歡豌
豆。於是，小明把午餐倒進超飳機，按了
一下超飳機上的按鈕。經過超飳機的處理，
豌豆就被去掉了。但是，超飳機只能去掉
一種食物。太挑食可不行。

(There are peas in Xiaoming's lunch, but he does not like peas. So Xiaoming poured his meal into 超飳機, and pressed the button on 超飳機. After 超飳機 processed the meal, there are no more peas in it. But 超飳機 can only remove one food at a time. Children should be too picky on the food.)

Appendix C. Statistical model structures

```
Maximal mixed-effects model = glmer (response ~ transparency*regularity*session +
                                     (1+transparency*regularity*session | item) +
                                     (1+transparency*regularity*session | subject),
                                     data = data,
                                     family = binomial,
                                     control = glmerControl (optimizer="bobyqa"))

Final model = glmer (response ~ transparency*regularity*session +
                     (1 | item) +
                     (1+regularity| subject),
                     data = data,
                     family = binomial,
                     control = glmerControl (optimizer="bobyqa"))
```