Theory and Practice of Second Language Acquisition
vol. 10 (2), 2024, pp. 1/22
https://doi.org/10.31261/TAPSLA.16086

Haotong Zhao
Beijing Foreign Studies University, China
https://orcid.org/0009-0005-1602-4685

The Role of Phonesthemes in EFL Learners' Word Acquisition

Abstract

Phonesthesia is one of the counterexamples of the arbitrariness of human languages. Although an individual word still appears arbitrary, a group of words bearing the same sound pattern might share similar meaning. This study investigated phonesthesia as a route to improving word acquisition. First, by comparing the guessing accuracy of phonesthemic versus prosaic words, we found phonesthemic words were significantly guessed better, suggesting EFL learners' sensitivity to English phonesthemes prior to explicit instruction. Phonesthemes provide hints for learners trying to ascertain the meaning of an unknown word. However, according to the participants' self-report, such sensitivity seems to lie somewhere in between pure clang association and genuine phonesthemic association. In another experiment, we examined the effect of phonesthemic knowledge, induced by explicit instruction, on the retention of phonesthemic words. The experimental group was taught 12 English phonesthemes while the control group was not. It was revealed that learners equipped with phonesthemic knowledge could retain phonesthemic words significantly better over an interval. We argue this is due to an extension of the lexical network in learners' minds. The second experiment further verified the facilitative role of phonesthemes in word retention. It is recommended that EFL teachers raise learners' attention to phonesthesia, which may potentially assist in word learning.

Keywords: phonestheme; sound symbolism; word learning; word retention

Sound symbolism is an exception to the long-held belief about the arbitrariness of linguistic signs. In recent years, de Saussure's (1983) proposition that "the bond between the signifier and the signified is arbitrary" (p. 62) has been constantly questioned. In the lexicons of many natural languages, there are words whose meanings are at least partially connected with their phonological

forms. Phonestheme is one example, which has been found in English (Firth, 1930), Swedish (Abelin, 1999), and Austronesian languages (Blust, 2003).

Phonestheme is operationally defined as the association of a sound pattern with a particular meaning. Monomorphemic words sometimes show form-and-meaning correspondences beyond the level predicted by chance (Bergen, 2004). For example, in English, word-initial *gl-* tends to occur in words related to light or vision such as *glitter*, *glimmer*, *glisten*, *gleam*, *glory*, *glow*, *glare*, *glance*, *gloss*, etc. Such sound-meaning pairings only have a predictive, not decisive effect on the meaning of a word and therefore are traditionally not considered compositional morphemes. Unlike morphemes whose meanings can be attested in every word containing them, phonesthemes can only be found in a limited group of words. *Glue*, *globe*, and *glove* share the consonant cluster /gl-/ as well, but are in no sense related to light or vision.

Empirical studies on English phonesthemes have predominantly focused on their psychological reality in native speakers. Using neologisms as stimuli, research has demonstrated speakers would interpret and produce unknown words in the exact pattern predicted by the phonesthemes in their language (Hutchins, 1998; Mueller, 2017). This finding refuted to the argument that phonesthemes are mere coincidences in the lexicon, because they do participate in the coinage and interpretation of new words. Another branch of psycholinguistic studies looked into how phonesthemes are represented and processed in speakers' minds. Priming tasks revealed that phonesthemic words demand less processing burden than words that are only semantically or orthographically related, and this tendency is alike to that reported for morphemes (Bergen, 2004). This means phonesthemes participate in language processing just like morphemes.

With the advent of corpora, scholars are also able to validate old phonesthemes, identify new phonesthemes and attach meaning to them from a statistical perspective. The results of these studies agree with each other on general. For instance, most of them found words containing a common phonestheme indeed exhibit greater relatedness than words selected randomly from a corpus (Otis & Sagi, 2008). However, it should be noted that different studies applied different mathematical models ranging from mutual information (Drellishak, 2006) to pairwise semantic similarity (Abramova & Fernández, 2016), so they provide slightly different answers as to which phonesthemes are statistically significant. For example, *tw-* is a significant phonestheme according to the criterion of Gutierrez et al. (2016), but not to Pimentel et al. (2019). The number of verified phonesthemes also differs according to the strictness of the applied standard. The study of Abramova et al. (2013) only generated six statistically significant phonesthemes while Liu et al. (2018) generated 30.

The role of phonesthemes in word acquisition, then, is still an underexplored topic. Phonesthemes have been found to scaffold word learning of native English speakers. Natives are able to produce better definitions for phonesthemic words

compared with the prosaic counterparts and recognize the meanings of phonesthemic words better (Parault & Schwanenflugel, 2006). This effect can be enhanced by the assistance of useful context (Parault & Parkinson, 2008).

Similarly, Zohrabi et al. (2014) suggested Iranian EFL learners can match unknown phonesthemic English words with the correct Farsi rendering better, indicating the sensitivity to phonesthemes in non-natives as well. However, this study does not report to have controlled for the influence of clang association. Clang association refers to the clustering of words bearing the same sound (Meara, 1983). In the study, participants were shown a phonesthemic word and required to choose one Farsi definition out of three. Participants would tend to choose a definition which contains a word phonologically (and therefore also orthographically) similar to the target word.

Sensitivity to English phonesthemes *tr-* and *fl-* is also attested in French, Spanish, and Macedonian speakers (Mompean et al., 2020). This is probably because the participants' native languages also have sub-morphemic items which seem to be associated with a particular meaning (Mompean et al., 2020). It is left unresolved, then, whether Chinese EFL learners can perceive English phonesthemes as well, because phonesthemes are more often reported in Indo-European and Austronesian languages, while the evidence of their existence in Mandarin Chinese and other Sino-Tibetan languages is scarce.

In terms of the effect of instruction on phonesthemes on word acquisition, to the best of our knowledge, there is little research addressing this topic, although such instruction can be very beneficial to EFL learners. First, English phonesthemes are not anecdotal but ubiquitous. A quick examination of Hutchins's (1998) list of 145 phonesthemes can reveal that almost every word-initial consonant cluster had been proposed by at least one scholar as carrying meaning association. Also, many monomorphemic words of Germanic origin are phonesthemic (He, 2002). They are relatively difficult to learn due to their indivisibility; learners must memorize the form-meaning correspondence as a whole. Phonesthemes, then, "provide a 'hook' for word learners trying to ascertain the meaning of an unknown word" (Parault & Schwanenflugel, 2006, p. 344).

Second, English phonesthemes are more productive than we may expect. Although one phonestheme may show different degrees of productivity due to the differing criteria on what qualifies as associated meaning, and different phonesthemes are, of course, productive to different extents, studies agree in general that the majority of English phonesthemes can reach >33% productivity (e.g., Bowles, 1995). That is, among words with the phonesthemic sound symbols, more than one third of them also contain the corresponding associated meaning. If a looser criterion is employed, this percentage can even rise to a striking 64% (Williams, 2021). Therefore, if phonesthemes were proved

conducive to word acquisition, EFL learners can apply phonesthemic knowledge to a fair number of words.

Based on the research gaps, the current study poses the following questions:

1. Are Chinese EFL learners sensitive to the sound-meaning connection of English phonesthemes without instruction?
2. After explicit instruction on English phonesthemes, can phonesthemic knowledge aid in EFL learners' word retention?

Methodology

Two experiments were devised to approach research question 1 and 2 respectively. In Experiment 1, by comparing the probability of participants choosing the correct definition for a phonesthemic word versus a non-phonesthemic word, we tried to assess EFL learners' sensitivity to English phonesthemes. Experiment 2 explored whether phonesthemic knowledge, induced by explicit instruction on English phonesthemes, improves retention of phonesthemic words. We taught the experimental group 12 English phonesthemes and tested whether they retained phonesthemic words better in memory over an interval.

Experiment 1

Participants

Participants in Experiment 1 were 49 non-English major undergraduates (21 males and 28 females) whose mother tongue was Mandarin Chinese. All of them had passed the College English Test Band 6 (CET-6), which required a minimal vocabulary size of 5,500 (*Syllabus*, 2006). Having learned English for at least 12 years, the participants were deemed advanced learners. The rationale for selecting advanced learners was that since phonesthesia is a conventional mapping due to recurrent sound patterns in the lexicon (Cuskley & Kirby, 2013), advanced learners who have a larger vocabulary tend to have more exposure to the patterns, making them more likely to recognize English phonesthemes, if it were possible. Moreover, to exclude the possibility that a participant was already aware of the form-and-meaning connection in phonesthemes, a brief interview conducted after the guessing test (see Experiment 1, Procedure & Instruments) revealed none of them had heard about the term *phonestheme* or received any instruction on it. Participation was voluntary, and each participant received a small compensation.

Stimuli

Stimuli in Experiment 1 consisted of 60 words, among which 24 were phonesthemic target words, 24 were non-phonesthemic target words, and 12 were fillers.

Drawn from Shipley's (1955) *Dictionary of Early English* and MacKay's (1879) *The Lost Beauties of the English Language*, all target words were obsolete English words that had "dropped from general use" (Shiply, 1955, p. ix). Obsolete words were employed to rule out any possibility that the participants had prior knowledge of them while at the same time ensuring authenticity.

To retrieve a collection of phonesthemic words from the dictionaries, previous research which statistically validated the reality of particular phonesthemes was reviewed (see the first section of the article). Twelve phonesthemes that have statistical evidence from at least four studies were taken into our scope. We restricted the target phonesthemes to those that occupy the initial position in words because they had more statistical and psychological evidence than rime ones. Then the dictionaries were surveyed in search of words which began with one of the 12 phonesthemes and carried the corresponding associated meaning. Only monomorphemic words with one to two syllables and four to seven letters were chosen. Each phonestheme appeared twice in the set of target words. The 24 phonesthemic target words are presented in Table 1.

Table 1

List of Phonesthemic Target Words in Experiment 1

Phonestheme	Associated meaning (Hutchins, 1998)	Target word and definition (MacKay, 1879; Shipley, 1955)
bl-	"blow, swell, or inflate; be round, swollen, or globular in shape"	*bleb*: a bubble of air in water or glass
		blore: a violent blowing
cl-	"two things coming together often producing a noise"	*clenge*: clangor
		clicket: a latch of a door
dr-	"having a languid, listless quality"	*drumble*: an inert or sluggish fellow
		dretch: to linger
fl-	"rhythmic motion; characteristic motion of liquid or gases"	*flabel*: to fan
		fleam: to flow
gl-	"light or vision, something visually salient"	*glaik*: a flash of strong light
		gledy: glowing hot

Phonestheme	Associated meaning (Hutchins, 1998)	Target word and definition (MacKay, 1879; Shipley, 1955)
gr-	"deep-toned, complaining, or threatening noise"	*grame*: to cause anger and grief
		grot: to bewail
scr-/skr-	"unpleasant sound"	*scrannel*: harsh, unmelodious
		scriek: to scream
sn-	"nose or breathing"	*snirt*: to laugh in a suppressed manner
		snite: to wipe the nose
st-	"firm, upright, regular, or powerful"	*staddle*: a foundation
		stith: unyielding, strong
str-	"forceful action in a line; something linear"	*streck*: straight
		streek: to stretch
sw-	"oscillate, undulate, or move rhythmically to and fro"	*swither*: to be uncertain or undecided
		swang: to swing
tw-	"turn, distort, entangle, oscillate"	*twirk*: to twirl
		twage: to tweak

The 24 non-phonesthemic target words were randomly extracted from the two aforementioned dictionaries. They were similar in length to their phonesthemic counterparts, but did not contain any sound symbol that was proposed to be phonesthemic.

The other 12 words served as fillers. Taken from *A Frequency Dictionary of Contemporary American English* (Davies & Gardner, 2010), they also had similar length to other stimuli and the frequency of the words' roots ranked 4,000th to 5,000th in American English. Words in this frequency range were thought to be mostly known by the participants, whose vocabulary size was required to reach 5,500 (*Syllabus*, 2006). The addition of familiar words aimed at enhancing the meaningfulness of the multiple choice test mentioned below, otherwise the participants might become reluctant to continue after making guesses about definitions of apparently unknown words.

Procedure and Instruments

Each participant first took a multiple choice guessing test to assess their sensitivity to the 12 target English phonesthemes. The 60 stimuli were presented in random order. Every word appeared along with four possible definitions, one being the correct answer and the rest three being randomly picked out from

Oxford English Dictionary (Oxford University Press, n.d.). The dummy options shared the same part of speech as but were not in any sense semantically connected to the intended answer. To preclude the potential influence of clang association (i.e., association of phonetically similar but semantically unrelated words, e.g., *reflect-effect*), if the four definitions had a word that shared the same onset consonant with the stimulus, it was replaced with a synonym or paraphrased. For example, the correct definition of *grame* was not presented as "to cause anger and *grief*" but "to cause anger and *bitterness*." Moreover, any words whose root was not one of the 5,000 most frequent words in *A Frequency Dictionary of Contemporary American English* were paraphrased with frequent words so that the participants would encounter fewer problems understanding the options. The definition of *drumble*, for instance, was changed from "an *inert* and *sluggish fellow*" to "a *slow* and *stupid person*." Participants were asked to choose the definition they reckoned to be the most appropriate. A sample of the test is given below:

grame
A. to express one's opinion
B. to be hidden
C. to cause anger and bitterness
D. to recover after illness

Participants were allowed to take as long as they want to finish the test because they are expected to utilize all their knowledge to infer the meaning of the unknown words. On average, the guessing test took around 30 minutes.

Participants were then asked to assess their prior knowledge of each stimulus against a scale created by Dale et al. (1986), where word knowledge is categorized into four levels ranging from totally no previous knowledge to a fully understanding of the word's meaning:

A. I never saw it before.
B. I have heard of it before but I don't know what it means.
C. I recognize it—it has something to do with…
D. I know it.

This was to ensure the absolute novelty of the targets to the participants. Since all of the target words are obsolete, it is very unlikely that a participant has encountered a word before or even know its meaning. Therefore, if any participant rated B or above for a target word, their answer to the corresponding question would be eliminated.

Lastly, every participant did a short interview. They were asked whether they had heard about the term *phonestheme* and whether they had received

explicit instruction on this phenomenon in their previous learning experience. If a participant had learned about English phonesthemes in school, it would be natural that they are sensitive to the phenomenon, and their response in the guessing test should be invalidated. Luckily for our purpose, none of them reported they were explicitly taught phonesthemes before.

Scoring and Data Aanalysis

Only the 48 test items based on (non-)phonesthemic target words were graded. For each correct response a participant received one point. Each participant's score was then divided by the number of valid test items to get the guessing accuracy. An independent sample *t*-test was conducted to compare the guessing accuracy of phonesthemic versus non-phonesthemic words.

Experiment 2

Participants

95 non-English-major undergraduates (25 males and 70 females) who had passed CET-6 took part in Experiment 2. None of them was also a participant in Experiment 1. Participants were randomly assigned to experimental group ($N = 44$) and control group ($N = 51$). An interview after the experiment showed they had no prior knowledge about English phonesthemes: They had not been instructed on this phenomenon, nor were they able to recognize the connection of phonesthemic patterns and their associated meaning by themselves. Every participant received a monetary reward for their participation.

Stimuli

Stimuli in Experiment 2 were a subset of the 24 phonesthemic target words in Experiment 1. From the 24 words, 12 were randomly selected to be the material for Experiment 2. Each word contained a unique phonestheme. The number of times each word being selected was counterbalanced among the participants. All 24 phonesthemic words were not used to relieve the burden of memorizing too many new words in the learning stage.

Procedure and Instruments

Experiment 2 consisted of five sections: instruction stage, learning stage, immediate posttest, interview, and delayed posttest. The experimental group went through all five sections, while the control group skipped the instruction stage and began from the learning stage.

In the instruction stage, the experiment group were introduced the phenomenon of phonesthesia and then presented a list of the 12 chosen phonesthemes, their associated meaning, some examples and a few counterexamples. A sample was presented in Table 2.

Table 2

Sample of the Material in the Instruction Stage of Experiment 2

Sound	Associated meaning	Examples	Counterexamples
gr-	complaining or threatening; negative emotion	grave, greed, grief, gross, grouse, growl, grudge, gruff grumble, grumpy, grunt	grape, great, group, grow

This was accompanied by the researcher's oral instruction while necessary. Participants were not required to memorize the associated phonesthemic meaning word for word. Instead, they were told to familiarize themselves with the example words so that they could deduce the associated meaning from the phonesthemic words in their own vocabulary. This stage lasted for 70 to 80 minutes on average, and the participants were given a one-day rest to integrate phonesthemic knowledge into their existing lexical knowledge.

In the learning stage, every participant tried to memorize 12 phonesthemic words that contained the target sound symbols. First, participants saw a word list of the 12 target words. In the list were the words' spelling, definition and a sample sentence. The definitions were extracted from the obsolete English dictionaries (MacKay, 1879; Shipley, 1955) and the sample sentences were constructed by the researcher. Any clang association words and infrequent words were paraphrased (cf. Experiment 1, Procedure and Instruments). An example is given below:

bleb
n. a sphere of air in water or glass
When fish want to get rid of excess air, they make **blebs**.

The participants were given 20 minutes and allowed to use whatever strategy they prefer, but for the participants in the experimental group, the researcher required them to apply the phonesthemic knowledge they learned a day ago. Every participant was also asked to report their previous knowledge about the target words against the Dale et al. (1986) checklist. Like in Experiment 1, if a participant rated B or above for a target word, their response to the corresponding items in the following tests would be deemed invalid.

Right after the learning stage, participants took a cloze test which measured immediate learning effect. In the test, participants saw a word bank of the 12 target words and 16 sentences each with a word removed, and were required

to insert the target words into the suitable sentence. The sentences worked as "general context." That is, they "provide clues to the [word's] meaning, although the specific characteristics of the word remain undefined" (Beck et al., 1983, p. 179). As long as they remembered the word's meaning, participants should obtain enough hints from the sentences to select the correct item. Four sentences functioned as distractors; no word fit into the blanks in them. The test lasted for 12 minutes. Below are two sample test items:

At midnight I was waked up by the _____ of the alarm.
From here the river _____ east into the ocean.
Word bank: **clenge fleam**

The cloze test lasted for 12 minutes.

This was followed by a brief interview, where we first asked every participant whether they had received instruction on phonesthemes before. If a participant was taught this phenomenon previously, they would have an advantage compared with the participants who had received no instruction. Luckily, none of them had reported previous knowledge on English phonesthemes. And we also selected nine of the participants in the experimental group and inquired their "feeling" while learning phonesthemes and memorizing phonesthemic words. We asked how they felt while learning English phonesthemes, whether such knowledge is compatible with their existing lexical knowledge and whether phonesthemes were helpful in learning new words. This was to make clear the participants' train of thought and supplement the findings based on the quantitative test scores. Note that the participants of Experiment 1 and 2 are from the same linguistic and educational background, so the results from this interview were also used to interpret the findings of Experiment 1.

Finally, to measure how well participants retained the target words in their memory, the cloze test was conducted again after a delay. The length of the delay was seven to eight days ($M = 7.21$). The test format of the delayed posttest was identical to the immediate posttest, but the exact items were completely different. For example, in the immediate cloze, the item designed for *swither* "to hesitate" is "The cellphone rang. George _____, not knowing whether to answer it." But in the delayed cloze the item designed for *swither* is "Mary _____a lot as to whether she should enter the arts or physics."

Scoring and Data Analysis

In the cloze tests, participants were given one point if they could find the correct lexeme for the suitable context; spelling and grammatical mistakes were excused. Their scores were divided by the number of valid test items to get

the accuracy. Finally, the experimental group's accuracy in the posttests was compared with the control group with an independent sample *t*-test.

Pilot Studies

A few pilot studies were conducted to ensure the validity of the instruments. First, some sentences in the material were written by the researcher, a native Mandarin Chinese speaker. Therefore, a native English speaker from the UK was employed to ensure the sentences were not awkwardly constructed. Second, we invited an extra five Chinese EFL learners of the same proficiency level as the participants to guarantee the time allocated to each task was sufficient for the participants to finish it. For example, 20 minutes was enough for the participants to memorize the meanings of 12 new words even through rote learning. Lastly, in terms of item design, the five EFL leaners also confirmed that (1) the 12 filler words in Experiment 1 were indeed familiar words to most of the participants. The addition of these words does enhance participants' confidence while taking the test. (2) In Experiment 2, participants could get full marks in the posttests when they have the words' definitions at hand, even with the presence of distractors.

Results

Experiment 1: Chinese EFL Learners' Sensitivity to English Phonesthemes

Experiment 1 intended to examine Chinese EFL learners' sensitivity to English phonesthemes. In the guessing test, we collected 24*49 = 1,176 responses for phonesthemic words and also 1,176 for non-phonesthemic words. 105 responses for phonesthemic words and 83 responses for non-phonesthemic words were deemed invalid according to the self-reports in the Dale et al. (1986) checklist. Therefore, 1,071 responses for phonesthemic words and 1,093 responses for non-phonesthemic words were taken into statistical analysis.

An independent *t*-test revealed participants were able to choose the correct definition of a phonesthemic word (M = .341, SD = .110) at a significantly higher rate than a non-phonesthemic word (M = .264, SD = .094), $t(96) = 3.716$, $p < .001$. The size of the effect is medium, Cohen's d = .75, 95% CI [.34, 1.16]. This finding supports the claim that participants can make use of the latent phonesthemic information to infer the meaning of an unknown word, indicat-

ing Chinese EFL learners can notice the sound-meaning connection of English phonesthemes without instruction.

The definition of a non-phonesthemic obsolete word was almost guessed randomly, as a one-tailed *t*-test showed no statistical difference between the observed accuracy and the likelihood of participants making the correct guess solely by luck (25%), $t(48) = 1.053$, $p > .05$. It suggests there is nothing within a non-phonesthemic word that can help a participant detect its meaning, making it more reasonable to attribute the high guessing accuracy of phonesthemic words to the existence of phonestheme.

Further Precluding the Influence of Clang Association

As stated above, an obsolete word might activate a phonologically (and therefore orthographically) similar present-day word in learners' mental lexicon, a phenomenon known as clang association. In Experiment 1, we had partly controlled for its effect by replacing any word in the four options which shared the same onset phonestheme as the target word with a dissimilar synonym. But there was still a possibility that participants' choice of the right option was based on mediated priming (Parault & Schwagenflugel, 2006). For example, one might have chosen the correct definition of *streck* "having no curve" because of its formal resemblance to the present-day word *stretch*, not because they had discovered anything special about *streck* itself. As it is technically impractical to manipulate participants' train of thought, the clang association strength of phonesthemic and non-phonesthemic words was then compared with Orthographic Levenshtein Distance 20 (OLD20).

OLD20 represents the average Levenshtein distance between a particular word and its 20 closest orthographic neighbors in the English lexicon. The minimum being 1, a lower OLD20 value indicates a denser neighborhood (Yarkoni et al., 2008), making the word more prone to clang association. The OLD20 scores of each obsolete word in the guessing test were examined using the Wuggy program (Keuleers & Brysbaert, 2010). An independent sample *t*-test showed no effect of word type on OLD20 value, $t(38) = .642$, $p > .05$. The phonesthemic words ($M = 1.890$, $SD = .239$) and non-phonesthemic words ($M = 1.831$, $SD = .337$) have an orthographic neighborhood of similar density. If the results of the guessing test were to be influenced by clang association, it is reasonable to expect a similar level of mediated priming effect across all target words. Clang association is therefore unlikely to have contributed to the differing guessing accuracy of phonesthemic and non-phonesthemic words.

Experiment 2: The Facilitative Role of Phonesthemes in Word Retention

Experiment 2 explores the facilitative role of phonesthemes in word retention. Prior to statistical analysis, participants' rankings of words based on the Dale et al. (1986) checklist were used to eliminate target words for which participants had previous knowledge. This accounted for 8.9% of the data.

In the immediate posttest, an independent sample *t*-test revealed the experimental (M = .859, SD = .139) and control (M = .857, SD = .128) groups obtained similar scores, $t(93) = .080$, $p > .05$. Both groups answered 10–11 of the questions correctly. With or without the instruction on phonestheme, all participants were able to reach a nearly perfect score in the immediate cloze test.

The results from this test revealed three facts. One, the mean scores of the experimental and control groups are similar, indicating there is no inter-group difference in the participants' memorizing ability. Both groups were able to memorize the target words equally well within the given time, which verifies the randomness of group assignment. Participants in the experimental group per se were no better word learners than the control group. Two, the standard deviations of both groups' scores are low. It means intra-group difference in memorizing ability is also negligible. No participant performed conspicuously poorly in the learning stage. This has dispelled a previous concern that an online experiment might not be taken seriously. Since all participants are similarly advanced English learners, this finding proves they had all done their best trying to memorize the target words. Lastly, both groups reached a relatively high accuracy, which shows the 20-minute learning time is sufficient enough to allow every participant, regardless of their learning strategy, to memorize the target words almost perfectly. Overall, results from the immediate posttest are proven to be a valid benchmark against which the degree of attrition can be measured in the delayed posttest.

In the delayed cloze, by contrast, the experimental group performed significantly better than the control group. While the control group (M = .457, SD = .255) experienced a sharp memory attrition over the interval, leading their test score to a 46.7% decrease, the experimental group (M = .597, SD = .245) managed to retain their memory better, their score decreasing by only 30.5%, $t(93) = 2.726$, $p < .05$. The size of the effect of instruction on delayed cloze test results is medium, Cohen's d = .56, 95% CI [.14, .97]. While the immediate cloze test proved both groups to have memorized the target words equally well, after the interval, the experimental group who had received instruction on phonestheme outperformed the control group by 34.7%. It lends evidence to the claim that phonesthemic knowledge can aid in EFL learners' word acquisition by promoting retention. The target words are significantly better retained in the memory of the participants in the experimental group.

Discussion

Non-arbitrary Form-to-meaning Link in Phonesthemes

The findings from Experiment 1 strongly support the claim that the link between form and meaning is not completely arbitrary (Dingemanse et al., 2015). Since the participants were only given decontextualized words to infer their meaning, it is only reasonable to attribute the phonesthemic words' high guessing accuracy to a feature intrinsic to the words themselves. In this experiment, we have tried to preclude the influence of every factor except for phonestheme which might guide the participants to a closer guess. The potential effect of clang association had been ruled out by two means: (1) Definition words that contain the same phonestheme as the stimuli were replaced with a synonym, and (2) the clang association strength was proven to be similar across all stimuli. Every target word, phonesthemic or not, should prompt a similar degree of clang association in the participants' mental lexicon. Thus, we argue clang association and orthographic neighboring cannot explain why participants guessed phonesthemic words significantly better, and that the effect of phonestheme cannot be reduced to an entirely sound- or spelling-based association. Phonesthesia should be treated as an independent phenomenon instead of a linguistic coincidence or a byproduct of clang association.

Sensitivity to English Phonesthemes

Experiment 1 also provides preliminary evidence for the psychological reality of English phonesthemes in Chinese EFL learners. As the participants scored higher on phonesthemic words on average, it proves they were able to perceive this phenomenon without prior instruction. This finding echoes previous experiments which attested sensitivity to English phonesthemes in native English speakers (Bergen, 2004) as well as French, Spanish, Macedonian (Mompean et al., 2020) and Korean EFL learners (Mueller, 2017). The present experiment is the first to find spontaneous recognition of English phonesthemes in participants whose mother tongue is a Sino-Tibetan language, where phonesthesia has not been reported. That L1 plays only a minimal role in affecting language learners' sensitivity to phonestheme seems to corroborate the claim that phonesthesia is conventional (Cuskley & Kirby, 2013). The form-to-meaning mappings are more likely to be a result of recurrent phonetic patterns in the lexicon of a particular language. Learners from whatever language background,

if fluent enough in English, are all able to recognize this phenomenon through repeated encounters with phonesthemic words.

However, whether such implicit sensitivity can be categorized as real phonesthemic knowledge remains a question. In the interviews of Experiment 2, a few participants reported they could recognize the phonetic and semantic similarities between some of the phonesthemic words. For example, one said they were aware that "*glitter*, *glisten* and *glimmer* have both similar appearance and similar meaning," but other *gl-* phonesthemic words like *glow* and *gleam* are not incorporated in the system, although *glow* and *gleam* were not new words to them. Sensitivity to the form-and-meaning link between some other words was also reported, but none of those could qualify as bona fide phonesthemic knowledge in that such sensitivity is only limited to two or three items. It seems that the form-and-meaning similarity between some of the phonesthemic words (when they are similar *enough*) is identifiable, but learners cannot pin down the exact word component (i.e., phonestheme) which contributes to such similarity or generalize this tendency to other words with the same component. Such sensitivity is, at best, an "implicit evaluation" of English phonesthemes rather than a "conscious elaboration" of the form-and-meaning link (Deconinck et al., 2010, p. 7).

In light of the self-reports of participants in Experiment 2, we argue EFL learners cannot form a full-fledged recognition of English phonesthemes. What is more likely to have happened, then, is that EFL learners can find a lexical association which lies somewhere in between pure clang association and genuine phonesthemic association. This association is not entirely form-based like *reflect-effect*, but it is also not large enough to include multiple (more than three in terms of our inspection) words that have the form-meaning correspondence. The question of whether it qualifies as phonesthemic knowledge needs further psycholinguistic research on whether EFL learners perceive English phonesthemes in the same way as native speakers. If EFL learners could show similar behaviors in psycholinguistic tasks (e.g., priming task) as natives, we can be more confident that phonesthesia is a psychologically real phenomenon in EFL learners' minds as well.

Scaffolding Effect of Phonesthemic Knowledge on Word Meaning Retention

Overall, Experiment 1 showed that Chinese EFL learners are able to mobilize phonesthemic information when encountering an unknown word. Phonesthemes, then, may be a useful lexical property from which EFL learners can draw hints to infer possible word meaning. As Hulstijn (2001) argues, "If a new word appears to the learner as having a form unrelated to its meaning,

it will need more attention and mental elaboration than if it has a transparent appearance" (p. 262). Therefore, phonesthemic words, having form-to-meaning transparency, might require less attention and mental burden of learners trying to memorize them. This contributes to more efficient learning and more endurable memory. Even in the absence of any other type of clue, phonesthemes still provide a "hook" for learners trying to deduce the meaning of an unknown word (Parault & Schwanenflugel, 2006, p. 344). In real learning situations where no options are given, the chance of guessing the meaning of a word correctly is very slim. But with phonesthemic knowledge in mind, learners can narrow down the possible meaning out of an infinite number of options.

The results from the delayed cloze test in Experiment 2 lend further support to the potentially facilitative role of phonesthemic knowledge in word retention. Participants in the experimental group, equipped with phonesthemic knowledge, underwent a slighter degree of memory attrition over the interval and, as a result, performed significantly better in the delayed test compared with the control group.

In terms of how phonesthemic knowledge improves word retention, we argue this is due to an extension of the lexical network in learners' mental lexicon. The interviews in Experiment 2 showed that the participants, though being advanced EFL learners, could only establish a word network of a few phonesthemic words, if it were possible at all. Explicit instruction on phonesthemic knowledge, then, helps to extend such a small, implicit lexical network to a larger network which includes more phonesthemic words in the learners' lexicon. Learners will be explicitly aware that the meaning of the words within a particular phonesthemic network applies to other members of the set as well, which reinforces the phonesthemic connection (Benczes, 2020). It will then become easier to capture new stimuli and incorporate them into the network. Words joined in the set, connected with other words, would be less susceptible to decay than words stored separately because they can also be retrieved through other words within the same network.

This finding resonates with processing models. According to the Levels of Processing (LOP) model, deep processing (e.g., semantic processing) is more immune to rapid memory attrition than shallow processing (e.g., phonetic and orthographic processing) (Craik & Lockhart, 1972). Since phonesthesia involves analysis of a word's meaning connection with other lexical items, it fosters a deeper-level information processing than rote learning and therefore makes the word more likely to be retained in memory. The Transfer Appropriate Processing (TAP) model differs from the LOP model in that it proposes semantic processing is not superior to structural processing in nature; instead, "the value of particular acquisition activities must be defined relative to particular goals and purposes" (Morris et al., 1977, p. 528). In other words, learning strategies which emphasize semantic elaboration result in better recollection

of meaning, while learning strategies which emphasize structural elaboration result in better recollection of form. According to the TAP model, phonesthemic knowledge prompts learners to evaluate the form as well as the meaning of a new word. Words memorized in this way should be better retained in terms of both form and meaning. However, it should be noted that the cloze test in Experiment 2 only assessed meaning recall, while form recall was not investigated in our study. So, at this stage it is only safe to conclude that phonesthemic knowledge improves meaning retention of unknown words based on our results. Whether it also facilitates form retention remains uncertain. This is exactly the question proposed by the Type of Processing Resource Allocation (TOPRA) model. The TOPRA model argues the processing resources available to a learner are limited, so semantic elaboration *improves* learning of semantic properties of words but *inhibits* learning of structural properties of words, and vice versa (Barcroft, 2002). As Experiment 2 showed phonesthemic knowledge improves meaning retention of new words, it could be at the expense of form retention. That is, learners equipped with phonesthemic knowledge might allocate too much attention to word meaning so that they would remember word forms less well than normal learners. Yet the opposite is believed to be true. Applying phonesthemic knowledge, learners do not cope with a word's form and meaning separately. Instead, they are encouraged to map them together. It is not an exclusively form-oriented or an exclusively meaning-oriented learning strategy. There would be at least weaker dispersion of learners' attentional resources. Future research can explore whether phonesthemic knowledge really induces learners to deal with form and meaning at the same time or it targets word meaning more.

Phonestheme-motivated Teaching Approach

Given that phonesthemes are proved facilitative in word meaning retention, the pedagogical implications of such form-meaning relationship are discussed. we propose that a phonestheme-motivated teaching approach should not force learners to remember the associated meaning of each phonestheme by rote. Instead, language teachers are recommended to raise learners' awareness to the presence of phonesthemes in English lexicon and help them transfer the "implicit evaluation" of phonesthemes to a "conscious elaboration" of such form-to-meaning mapping (Deconinck et al., 2010, p. 7), as was done in experiment 2. This is not only because rote learning isolated fragments of knowledge from context (Mayer, 2002) and is thus argued to be inferior to other supposedly more efficient word learning strategies (Nakamura, 2000). More importantly, phonesthemes play a predictive, rather than decisive, role in the meaning construction of a word. It is not rare that a word contains a phonestheme but does

not carry the proposed associated meaning. *Great*, for example, is not semantically connected to the phonesthemic meaning of *gr-* "deep-toned, complaining, or threatening noise." It even lies at the exact opposite end of the complaining spectrum. Also, how the associated phonesthemic meaning is integrated in a word's lexical meaning is subjective, depending on language users' interpretation. It is obvious that the meaning of *gl-* "light and vision" is perfectly integrated in the meaning of *glow* "give out light." But for *glass* the meaning of phonestheme *gl-* becomes more obscure. Therefore, learning phonesthemes by rote might cause learners to overemphasize the contribution of phonestheme to a word's meaning. An inductive teaching approach might be more efficient for phonesthemes. In other words, learners are trained to be more sensitive to the form-and-meaning similarities among words already in their vocabulary, and encouraged to derive a form-to-meaning mapping pattern based on extant lexical knowledge.

Morphological Status of Phonesthemes

Lastly, the question of whether phonesthemes should be treated as normal morphemes is of concern. Traditionally, morphemes are considered the smallest units carrying meaning. Although phonesthemes also involve a pairing of form and meaning, they differ substantially from morphemes. First, phonesthemes are non-compositional. Removal of the phonestheme from a word only results in a sequence of meaningless letters. It would be problematic to argue that *immer* contributes to the meaning of *glimmer*. Another difference is that morphemes are defined by contrast while phonesthemes by recurrent association (Blust, 2003). A single case is sufficient to identify an isolated morpheme. On the other hand, since phonesthemes can only be identified through recurrent phonetic patterns, multiple examples are necessary to support the existence of a phonestheme. Most scholars maintain phonesthemes are sub-morphemic components of a word (Blust, 2003). Yet some argue that a phonestheme is "a derivational morpheme in its own right" (Rhodes & Lawler, 1981, p. 325).

The present study does not intend to make any theoretical assertions. Nevertheless, it does show morphemes and phonesthemes exhibit similarities in terms of processing. We found phonesthemic knowledge improves advanced EFL learners' retention of phonesthemic words. This resonates with previous studies which demonstrated that morphological knowledge facilitates advanced EFL learners' acquisition of morphologically complex words (Sukying, 2020) and increases immature native speakers' ability to infer the meaning of an unknown complex word (Baumann et al., 2003). The instruction on both morphemes and phonesthemes is beneficial for the acquisition of the corresponding type of words. Moreover, there is also evidence that native English

speakers perceive and store phonesthemes in their lexicon in a manner much alike to normal morphemes (Bergen, 2004). It seems that phonesthemes also participate in language processing and language acquisition like morphemes, although they are non-compositional.

Conclusion

The present study generated two major findings. First, advanced Chinese EFL learners are generally sensitive to the phonesthemic (i.e., both phonetic and semantic) links between some of the English words bearing the sound symbol. They can also apply this tendency to an unknown word to infer its possible meaning. However, such sensitivity seems to lie somewhere in between clang association and genuine phonesthemic association, as the lexical networks established by learners in this way only include a few phonesthemic items. Second, phonesthemic knowledge, which can be instilled into learners through explicit instruction, significantly improves word retention over a one-week interval. This is possible because explicit instruction enlarges the phonesthemic networks and reinforces the connection between members in the networks. Therefore, although EFL learners might not recognize this linguistic feature systematically, phonesthesia indeed plays a critical role in word learning and is a useful resource for promoting word retention. Given the ubiquity of phonesthemic words in the English lexicon, language teachers are recommended to raise learners' awareness of this phenomenon and give learners explicit instruction on the form-to-meaning mapping between phonesthemes when appropriate. This could arguably promote the meaning retention of unknown phonesthemic words significantly. Also, in teaching phonesthemes, it may be more efficient to encourage learners to derive the form-to-meaning mappings on their own instead of forcing learners to memorize a list of English phonesthemes by rote.

One limitation of the present study is that we did not investigate how English phonesthemes are processed and stored in Chinese EFL learners' minds, and therefore we could not confirm whether the sensitivity to English phonesthemes exhibited in Experiment 1 qualifies as real phonesthemic knowledge. It is only certain that the phonesthemic links between part of the phonesthemic words in the learners' lexicon are spontaneously identifiable. Future studies could take a psycholinguistic perspective. If Chinese EFL learners are found to process English phonesthemes in a pattern similar to native speakers, it will be evidence that EFL learners' recognition of English phonestheme is systematic, or at least native-level.

Another limitation is that in Experiment 2, the posttest assessed meaning retention, while form retention was not examined. According to the TOPRA model, when a learning strategy allocates learners' attention to the semantic aspect of the word, promoting meaning retention, it may at the same time hinder form retention because the total amount of mental resource available to a learner is limited. Although we argue phonesthemic knowledge is unlikely to impair form retention since it targets form and meaning holistically, such possibility cannot be completely wiped out. Future studies could apply a test format which examines form as well as meaning retention, and see if phonestheme knowledge also assists in learners' form retention of new phonesthemic words.

References

Abelin, Å. (1999). *Studies in sound symbolism* [Unpublished doctoral dissertation]. Göteborg University.

Abramova, E., & Fernández, R. (2016). Questioning arbitrariness in language: A data-driving study of conventional iconicity. *Proceedings of NAACL-HLT 2016*, 343–352.

Abramova, E., Fernández, R., & Sangati, F. (2013). Automatic labeling of phonesthemic senses. *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*, 1696–1701.

Barcroft, J. (2002). Semantic and structural elaboration in L2 lexical acquisition. *Language Learning, 52*(2), 323–363.

Baumann, J. F., Edwards, E. C., Boland, E. M., Olejnik, S., & Kame'enui, E. J. (2003). Vocabulary tricks: Effects of instruction in morphology and context on fifth-grade students' ability to derive and infer word meanings. *American Educational Research Journal, 40*(2), 447–494.

Beck, I. L., McKeown, M. G., McCaslin, E. S. (1983). Vocabulary development: All contexts are not created equal. *The Elementary School Journal, 83*(3), 177–181.

Benczes, R. (2020). Sound symbolism and semantic change. *Dynamics of language changes: Looking within and across languages*, 253–264.

Bergen, B. K. (2004). The psychological reality of phonaesthemes. *Language, 80*(2), 290–311.

Blust, R. (2003). The phonestheme ŋ- in Austronesian languages. *Oceanic Linguistics, 42*(1), 187–212.

Bowles, H. (1995). The semantic properties of the phonaestheme. *Studi italiani di linguistica teorica ed applicata, 24*(1), 91–106.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671–684.

Cuskley, C., & Kirby, S. (2013). Synesthesia, cross-modality and language evolution. In J. Simner & E. Hubbard (Eds.), *Oxford Handbook of Synesthesia*. Oxford University Press.

Dale, E., O'Rourke, J., & Barbe, W. B. (1986). *Vocabulary building: A process approach*. Zaner-Bloser.

Davies, M., & Gardner, D. (2010). *A frequency dictionary of contemporary American English: Word sketches, collocates, and thematic lists*. Routledge.

Deconinck, J., Boers, F., & Eyckmans, J. (2010). Helping learners engage with L2 words: The form-meaning fit. *Aila Review, 23*(1), 95–114.

de Saussure, F. (1983). *Course in general linguistics*. Columbia University Press.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences, 19*(10), 603–615.

Drellishak, S. (2006). *Statistical techniques for detecting and validating phonesthemes* [Master's thesis, University of Washington].

Gutiérrez, E. D., Levy, R., & Bergen, B. K. (2016). Finding non-arbitrary form-meaning systematicity using string-metric learning for kernel regression. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, 2379–2388.

He, C. (2002). Phonosemantics: The study of sound and sense. *Foreign Language Teaching and Research (Bimonthly), 34*(1), 22–29.

Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and Second Language Instruction*, 258–286. Cambridge University Press.

Hutchins, S. S. (1998). *The psychological reality, variability, and compositionality of English phonesthemes* [Doctoral dissertation, Emory University].

Keuleers, M., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods, 42*(3), 627–633.

Liu, N. F., Levow, G., & Smith, N. A. (2018). Discovering phonesthemes with sparse regularization. *Proceedings of the Second Workshop on Subword/Character Level Models*, 49–54.

MacKay, C. (1879). *The lost beauties of the English language*. Chatton & Windus.

Meara, P. (1983). Word associations in a foreign language. *Nottingham Linguistics Circular, 11*(2), 29–38.

Mompean, J. A., Fregier, A., & Valenzuela J. (2020). Iconicity and systematicity in phonaesthemes: A cross-linguistic study. *Cognitive Linguistics, 31*(3), 515–548.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*, 519–533.

Mueller, J. (2017). *An examination of the influence of age on L2 acquisition of English sound-symbolic patterns* [Doctoral dissertation, University of Maryland].

Nakamura, T. (2000). *The use of vocabulary learning strategies: The case of Japanese EFL learners in two different learning environments* [Unpublished doctoral dissertation]. University of Essex.

Otis, K., & Sagi, E. (2008). Phonaesthemes: A corpus-based analysis. *Proceedings of the 30th Annual Meeting of the Cognitive Science Society*, 65–70.

Oxford University Press. (n.d.). *Oxford English Dictionary.* https://www.oed.com.

Parault, S. J., & Parkinson, M. (2008). Sound symbolic word learning in the middle grades. *Contemporary Educational Psychology, 33*, 647–671.

Parault, S. J., & Schwanenflugel, P. J. (2006). Sound-symbolism: A piece in the puzzle of word learning. *Journal of Psycholinguistic Research, 35*(4), 329–351.

Pimentel, T., McCarthy, A. D., Blasi, D. E., Roark, B., & Cotterell, R. (2019). Meaning to form: Measuring systematicity as information. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 1751–1764.

Rhodes, R., & Lawler, J. M. (1981). Athematic metaphors. *Chicago Linguistic Society, 17*, 318–342.

Shipley, J. T. (1955). *Dictionary of early English*. Philosophical Library.

Sukying, A. (2020). Word knowledge through morphological awareness in EFL learners. *TESOL International Journal, 15*(1), 74–85.

Syllabus for CET-6. (2006). Shanghai Foreign Language Education Press.

Williams, C. (2021). From *slifor* to *Slytherin*: The relationship between word form and meaning. In S. Morgan & J. Savage (Eds.), *Research Notes 82*. Cambridge University Press & Assessment.

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's *N*: A new measure of orthographic similarity. *Psychonomic Bulletin & Review, 15*(5), 971–979.

Zohrabi, M., Sabouri, H., & Peimanfar, S. (2014). Intermediate level learners' vocabulary guessing and recognizing through sound symbolism. *International Journal of Applied Linguistics & English Literature, 3*(6), 42–52.