This paper should be cited as:
Liu, J., Gu, J., Feng, C., Shi, W., Biemann, C., & Li, X. (in press). *Scientific Studies of Reading*.

Cross-Modal Impact of Recent Word Encountering Experience

Jiayu Liu[1,2], Junjuan Gu[3], Chen Feng[1], Weiting Shi[1,2], Chris Biemann[4], and Xingshan Li[1]

[1]CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, China
[2]Department of Psychology, University of Chinese Academy of Sciences, China
[3]School of Psychology and Mental Health, North China University of Science and Technology
[4]Department of Informatics, Universität Hamburg, Germany

Author Note

This work was jointly funded by the National Natural Science Foundation of China (NSFC) and the German Research Foundation (DFG) in Project Crossmodal Learning, NSFC 62061136001 /DFG TRR-169. This research was supported by a grant from the National Natural Science Foundation of China (31970992).

Correspondence should be addressed to Xingshan Li, 16 Lincui Road, Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China. Email: lixs@psych.ac.cn.

Abstract

Purpose:

This study was designed to distinguish the degree of sharing of representations between different modalities by investigating whether a word encountering experience in one modality impacts word processing in another modality.

Method:

In three experiments, participants experienced some words frequently in the auditory modality (Experiment 1, sample size 30, mean age 23.4, 56.7% female, all participants were native Chinese speakers) or visual modality (Experiment 2, sample size 30, mean age 22.4 years, 63.3% female, all participants were native Chinese speakers) in the training session, and were tested whether the word encountering experience impacts the results of Chinese word segmentation in the visual modality in the test session. In Experiment 3 (sample size 30, mean age 24.6 years, 76.7% female, all participants were native Chinese speakers), we used a within-subjects design, in which each participant received both auditory and visual training tasks.

Results:

The results of the three experiments showed that encountering a word frequently in a short period of time in the auditory modality or visual modality can affect word segmentation results in Chinese reading, with a recently experienced word being more likely to be segmented as a word. This effect was long-lasting, as it could still be observed after 7 days.

Conclusion:

The results suggest that the effect of a word encountering experience in listening can be transferred to reading. Thus, word encountering experiences should be stored at a location in the mind that is used for both listening and reading.

Cross-Modal Impact of Recent Word Encountering Experience

The frequency at which we encounter a word affects how easily we can process it, with frequently encountered words being processed faster than less frequently encountered ones (Rayner, 1998). This effect has been observed in different tasks such as lexical decision (Broadbent, 1967; Gardner et al., 1987; Rubenstein et al., 1970), naming (Forster & Chambers, 1973), and eye movement studies on natural reading (Inhoff & Rayner, 1986; Just & Carpenter, 1980; Rayner & Duffy, 1986). The effect has also been shown to be reliable in both alphabetic writing systems and logographic writing systems such as Chinese (Wei et al., 2013; Yan et al., 2006), and it has been observed within different language processing modalities such as reading (Inhoff & Rayner, 1986), listening (Dahan et al., 2001), speaking and writing (Bonin & Fayol, 2002).

Previous studies have shown that word frequency is one of the most important word properties affecting word processing. For example, some studies showed that word frequency can account for 40% variance of lexical decision time (Balota et al., 2007; Brysbaert et al., 2016; Ferrand et al., 2018). Even though these studies have shown that word frequencies are important, it is less clear how word frequency information as experienced in different modalities is represented in the mind and how it affects word processing; practically, when investigating word frequency effect in reading, it is unclear whether word frequency should only include visual word frequency or also the frequency of other modalities. In the present study, we investigated whether recent word encountering experiences in the auditory modality (listening) impact word processing in the visual modality (reading), and we also investigated whether encountering words in the auditory modality has the same effect on reading as encountering words in the visual modality (i.e. within reading).

Shared or Distinct Mental Lexicon

Different reading models have made different assumptions regarding whether lexical knowledge is shared by listening and reading. In some models, the lexical representation systems of the auditory and visual modalities are independent of each other (Grainger et al., 2003; Ellis & Young, 1988; Grainger & Ferrand, 1994; Harris & Coltheart, 1986; Morton, 1979; Patterson & Shewell, 1987). For example, the last version of the logogen model includes different logogen libraries (also known as lexicons) for visual lexical recognition, spoken lexical recognition, and speech production (Morton, 1979, 1982). In each lexicon, words are characterized as “logogens” and are recognized when perceptual evidence accumulates beyond the threshold of the logogen. Thus, such models assume that lexical knowledge in different modalities has separate representations. These models were supported by the finding that the priming effect was observed when the prime and target were within the same modality (such as a visual word primes a visual word) but not across modalities (Kempley & Morton, 1982). In these models, each word unit has some special mechanism to represent word frequency, so that word frequency information can influence word identification. For example, in the logogen model, each time the word is encountered, the corresponding logogen fires, and its threshold decreases by a small amount (Morton, 1969); therefore, a high-frequency word is identified more rapidly than a low-frequency word because the threshold is lower. According to these models that assume distinct mental lexicons, word frequency information in a modality is stored in the lexicon that is specific to that modality, and therefore, word frequency acquired in one modality cannot transfer to other modalities.

Another category of word recognition models, the triangle model, assumes that words are represented and processed through three interactively connected subnetworks: orthographic, phonologic, and semantic (Harm & Seidenberg, 2004; Seidenberg & McClelland, 1989). These subnetworks work jointly to achieve the meaning of the word. Semantic information can be activated through two pathways: a direct semantic pathway and a phonologically mediated pathway. In the direct semantic pathway, the activations of orthographic nodes (which encode input visual information) directly activate semantic nodes through the connections between orthographic nodes and semantic nodes. In the phonologically mediated pathway, the activations of orthographic nodes first propagate to phonologic nodes and subsequently to semantic nodes. These two pathways work simultaneously for alphabetic reading. The weights between nodes are adjusted using training datasets, so that the network can map specific patterns of input nodes onto specific patterns of output nodes. In the model's initial state, the states of connections between all nodes are defined as initial values. Each time a word is presented, the model uses an error-correcting algorithm to adjust the weights of connections between different forms of nodes and semantic nodes. After training, word information such as word frequency is represented as weights of links between nodes at different levels. There are more high-frequency words than low-frequency words in the training set so that high-frequency words have more influence on adjusting the weights.

It should be noted that the triangle model proposed by Harm and Seidenberg (2004) also incorporates attractor structures, which were created by including feedback connections to the semantic nodes via a set of “cleanup nodes.” To implement this attractor structure, all semantic nodes are connected to all cleanup nodes, and all cleanup nodes are all connected back to the semantic nodes; therefore, word frequency can be represented as weights of some or all the following links: between orthographic and semantic, between orthographic and phonologic nodes, between phonologic and semantic nodes, and between semantic and cleanup nodes. The triangle model can make different predictions when different assumptions are made regarding where word frequency information is stored. If word information is stored in the weight of the links that are shared to listening and reading (such as semantic to cleanup nodes), the model predicts a cross-model transfer. However, if the model assumes that word frequency information is stored in the weights of links that are used only by one modality (e.g., the orthographic to semantic pathway or phonologic to semantic pathway), no cross-modal transfer is expected.

In summary, different models have made different assumptions regarding how word frequency information is represented, and different predictions have been made regarding whether word information acquired from listening can be transferred to other modalities. While some models that assume shared mental lexicons predict a cross-modality transfer effect of word frequency, other models do not. Even for a single model such as the triangle model, different assumptions lead to different model predictions.

Interactions Between Orthography, Phonology and Semantics

Consistent with the assumptions of the triangle model (Harm & Seidenberg, 2004), previous studies have shown clear evidence that both the direct semantic pathway and the phonologically mediated pathway are used during reading. Some studies have shown that phonology is automatically activated during alphabetic word reading (Daneman & Stainton, 1991; Drewnowski & Healy, 1982; McCrutchen & Perfetti, 1982; Pollatsek et al., 1992; see Brysbaert, 2022, for a review). For example, Van Orden (1987) showed that readers make more errors in a semantic decision task for words with a homophone neighbor at different semantic categories. In that study, readers often mistakenly classify “rows” (whose homophone neighbor is “rose”) as a flower than the control word “robs”.

Similarly, some studies also showed evidence that orthographic information is activated during spoken language processing (Muneaux & Ziegler, 2004; Pattamadilok et al., 2008, 2010; Seidenberg & Tanenhaus, 1979; Tyler & Burnham, 2006; Ventura et al., 2004). In one line of research, participants were asked to judge whether two spoken words rhymed (e.g., McPherson, Ackerman, & Dykman, 1997; Zecker, 1991; Zecker, Tanenhaus, Alderman, & Siqueland, 1986; Rack, 1985; Donnenwerth-Nolan, Tanenhaus, & Seidenberg, 1981; Seidenberg & Tanenhaus, 1979). Response times were shorter when the rhymes of two spoken words are spelled similarly (e.g., “toast” and “roast”) than when they are spelled differently (e.g., “toast” and “ghost”). A similar orthographic consistency effect has also been reported in an ERP study using a semantic categorization task, in which participants do not need to spell the word (Pattamadilok et al., 2009). When compared to consistent words, they found increased difficulty (as reflected in increased N400 amplitudes) in lexical access to spoken words that have an inconsistent mapping between phonology and orthography. The findings from these studies suggest that orthographic information can influence spoken language processing.

Taken together, these studies show that phonology and orthography are not processed independently. Instead, the phonologic code is activated during reading, and orthographic code is activated during spoken language processing. These findings suggest that experiencing a word frequently in the auditory modality might impact word processing in visual modality.

Previous Studies on Cross-Modal Transfer of Word Encountering Experience

Research in the listening and reading fields has developed separately, and until recently, very few studies had investigated them jointly (Ferrand et al., 2018; Kligler & Gabay, 2023). The findings from these few studies can shed some light on whether recent encountering experience of a word impacts word processing in another modality. These studies have used different paradigms and showed that encountering a word in one modality can impact word processing in another modality for a longer time.

Some studies investigated whether word frequency acquired through production (e.g., speaking) can be transferred to comprehension (e.g., reading; Van Assche et al., 2016). In one of their experiments, Van Assche et al. (2016) found that training of picture naming (which is a spoken production task) accelerated the response speed of lexical decision (which is a reading task). Meanwhile, the training of lexical decision also accelerated the response speed of picture naming. Therefore, the study suggested the existence of a modality independent representation (lemma) that is shared between production and recognition. In the models that assume shared lexical representations, lexical knowledge such as word frequency, as part of lexical representation, may be shared across modalities to some extent. Even though Van Assche et al. (2016) showed that lexical representation may be shared by production and comprehension, it does not mean that word frequency information can be shared by listening and reading.

Previous studies also showed that experiencing a word can impact which meaning of an ambiguous word is perceived later (Rodd et al., 2013; 2016). For example, Rodd et al. (2013) let participants listen to a word embedded in sentences so that one ambiguous word has a specific meaning in the specific sentence context. Participants preferred to choose the primed word meaning after they listen to the ambiguous word out of context in an association task 20 minutes later. More importantly, later studies also showed that the long-term semantic priming effect is modality-general (Gilbert et al., 2018). The semantic priming effect was found from the auditory to the written modality, and vice versa. These studies show that experiencing a word can affect word interpretation and the following word interpretation in different modalities, and the effect can last for at least 20 minutes.

Some corpus studies showed strong correlations between spoken word frequency and visual word frequency. For example, Brown and Watson (1987) indicated that the correlation between visual word frequency and auditory word frequency was 0.70. We should note that a correlation between spoken word frequency and visual word frequency does not necessarily mean that the two are represented in the same locus in the mind. An alternative reason for high correlation may be that a word used frequently in spoken language may also be frequently used in reading because it refers to a highly common concept. Moreover, other corpus studies have shown that the average of visual word frequency and spoken word frequency can account for more variance of rated familiarity and lexical decision latency than visual word frequency alone (Brown & Watson, 1987; Pastizzo & Carbone, 2007), indicating that spoken word frequency may carry some additional information to predict the performance of visual word processing that is not included in visual word frequency.

Because word frequency is one of the most important word properties that affect reading (Zhao et al., 2018), many researchers have attempted to find the optimal word frequency measures. In studies of different languages such as Chinese, English, and French, a corpus constructed from subtitles is significantly better at predicting lexical decision tasks in vision than a corpus constructed based on written language only (Brysbaert & New, 2009; Cai & Brysbaert, 2010; New et al., 2007). From these phenomena, we can find that spoken word frequency might have some implications worth considering when building a lexical resource for visual word frequency.

Although these studies provide some hint that recently encountered words in the auditory modality may impact word processing in the visual modality (reading), the evidence is not direct. The present study was designed to directly investigate this question.

Word Processing in Chinese Reading

In the present study, we investigated whether a recent word encountering experience in the auditory modality impacts word processing in Chinese reading. The Chinese writing system is an orthographic writing system that is different from the alphabetic writing system in many dimensions (Li et al., 2022 ; Yao et al., 2021). There are more than 5,000 Chinese characters, and each Chinese character represents a syllable and usually represents a morpheme at the same time. One important cross-language difference is how information is transferred through different pathways. While alphabetic reading relies on both the direct semantic pathway and the phonologically mediated pathway, some models of Chinese reading assume that orthographic reading mainly relies on the former, because the latter is slow and not efficient enough (Perfetti et al., 2005). Perfetti et al. (2005) used a connectionist model to simulate single-character Chinese word reading, mainly focusing on how words are pronounced. They proposed that the links between orthographic nodes and phonologic nodes are thresholded ones: the corresponding phonological nodes of a character are activated only when the character is identified. In Chinese, a spoken syllable is mapped onto a whole character, and no part of the character corresponds to any subset of a syllable; therefore, Chinese readers cannot know the pronunciation of the character before the character is fully identified. As a result, the phonologically mediated pathway is much slower than the direct semantic route, and Chinese readers usually use the latter to access word meanings during reading. We should also note that there is strong evidence that phonological information is activated during Chinese reading (Chua, 1999; Tan & Perfetti, 1997; Xu et al., 1999). Some recent studies also suggest that phonetic radicals also contribute to Chinese word reading, and fMRI studies showed different brain activation patterns when phonetic radicals are located at different sides of the characters (Liu et al., 2020). However, there is no strong evidence that showed that Chinese readers could achieve word meaning through the phonologically mediated pathway (Tan & Perfetti, 1997; Zhou et al., 1999; Zhou & Marslen-Wilson, 1999; see Perfett et al., 2005 for a review). This cross-language difference might cause different results regarding whether a recent word encountering experience in the auditory modality affects word processing in visual modality.

Another unique property of Chinese is that there are no inter-word spaces or other physical cues to demarcate words in Chinese text. Therefore, Chinese readers cannot rely on low-level visual information to group contiguous characters into words (word segmentation). Even though lack of inter-word spaces, there is strong evidence that readers process words as a whole during Chinese reading, just as readers of alphabetic writing systems do (Cheng, 1981; Li et al., 2009, 2012; Mok, 2009; Shen

& Li, 2012). One line of research using the Reicher-Wheeler paradigm (Reicher, 1969; Wheeler, 1970) is that a word superiority effect was observed in Chinese reading, which showed that a Chinese character is reported more accurately when it is embedded within a word than in a non-word string (Cheng, 1981; Mok, 2009; Shen & Li, 2012). Mok (2009) further showed that the size of word superiority effect is larger when at least one constituent of Chinese compound words is opaque than when both constituents are transparent, suggesting that the opaque words may be more word-like than full transparent compound words. Moreover, another consequence caused by the lack of inter-word space in Chinese reading is that Chinese readers do not always agree on where the word boundaries are (Hoosain, 1991; Liu et al., 2013). A recent study shows that how Chinese readers segment words is affected by the attributes of the reader, the words, and the characters (Chen et al., 2018).

Without inter-word spaces, how do Chinese readers segment words? A recent computational model, Chinese reading model (CRM), provides a solution to this question (Li & Pollatsek, 2020; see also Li et al., 2009). According to CRM, all of the characters in the perceptual span are processed in parallel. All words constituted by these characters are activated and they compete. When one word wins the competition, the word is identified and it is segmented from the text at the same time. Therefore, CRM assumes that word segmentation and identification happen simultaneously in Chinese reading, without one happening earlier than the other.

In the present study, we detected the impact of encountering a word in the auditory modality on visual word processing using a paradigm that was caused by a lack of inter-word spaces property of Chinese reading. Because there are no spaces between words in Chinese text, word segmentation is sometimes ambiguous, meaning that readers can segment words in different ways. One kind of ambiguity involves overlapping ambiguous strings (OAS). For example, in the string ABC (e.g., 学生活), the first two characters AB can form a meaningful word (e.g., 学生, means "student"), and the last two characters BC of the string can also form a meaningful word (e.g., 生活, means "living"). Chinese readers need to determine whether the middle character belongs to the word on the left or right. To illustrate this in English, a word "unlockable" can be interpreted as either "unlock-able" or "un-lockable", and readers usually can only perceive one meaning during reading (Pollatsek, Drieghe, Stockall, & de Almeida, 2010). Previous studies showed that the results of segmentation are mainly influenced by two factors: position and word frequency (Huang et al., 2020, 2021; Ma et al., 2014). First, the word AB located on the left has a certain position advantage because of the left-to-right reading habit in Chinese. When other factors are equal, readers prefer to segment the OAS string ABC as AB-C. Second, segmentation is influenced by word frequency so that a high-frequency word constituted by the OAS string is more likely to be segmented as a word. For example, Ma et al. (2014) manipulated the word frequency of AB and BC in their experiment, and the results showed that Chinese readers were more likely to segment the string ABC as A-BC rather than AB-C when BC had a higher word frequency than AB. Therefore, this study shows that the segmentation of OAS is sensitive to the frequency of the constituted words. In the present study, we used this property of Chinese reading as a method for detecting the impact of recent word encountering experiences on visual word processing. We chose this method because the effect is reliable and has been replicated in many studies (Huang et al., 2020, 2021; Ma et al., 2014).

The Present Study

As reviewed above, different models made different assumptions regarding how word frequency information is represented in the mind and made different predictions regarding whether recent word encountering experiences in the auditory modality (listening) impact word processing in the visual modality (reading). Even for a single model, such as the triangle model, different predictions can be made when the model adapts different implementation assumptions. Therefore, understanding whether word encountering experience from one modality impacts word processing in another modality is theoretically important because it can be used to constrain the development of word processing models. Nevertheless, as previous experimental studies have not provided strong evidence to support any strong conclusion on this issue, the current study was designed to address the question regarding whether recent word frequency information acquired from the auditory modality (listening) can be transferred to the visual modality (reading) in Chinese.

In the present study, we used three experiments to investigate the transfer effects of word encountering experience in cross-modal tasks and within-modality tasks. In Experiment 1, participants were trained in the auditory modality in the training session and were tested in the visual modality in the test session. In Experiment 2, participants were trained in the visual modality in the training session and were tested in the visual modality in the test session. We attempted to distinguish the degree of sharing representations between different modalities by detecting whether encountering a word in the auditory modality affects visual word processing between Experiments 1 and 2. Different models make different predictions regarding the results of the current study. Models that assume a shared mental lexicon predict a cross-modal transfer of word encountering experience, while models that assume distinct mental lexicons predict no such effect. In the framework of the triangle model, if word encountering experience is stored as weights of the links that are only used in the phonologic pathway, a cross-modal transfer effect is not expected to be found because previous studies showed that mature Chinese readers do not usually use the phonologically mediated pathway. However, if word encountering experience in the auditory modality (listening) is stored either in the orthographic to semantic pathway or at the semantic level, a cross-modal transfer effect is expected.

Comparing the effect size of Experiments 1 and 2 can further help to distinguish some important aspects of the triangle model. If the word encountering experience in reading is mainly stored in the weights of links between orthographic nodes and semantic nodes, then a larger effect of training is expected in the visual modality. However, if the word encountering experience is mainly stored at the semantic level, the effect size of both experiments should be comparable.

Experiment 3 was designed to replicate the findings of Experiments 1 and 2, and to exclude some other possible explanations of the results. In Experiments 1 and 2, the effects of word encountering experiences in different modalities were compared between participants. In Experiment 3, participants were trained in both auditory modality and visual modality (for different items) with a within-participant design, which could help to reduce errors associated with individual differences. Moreover, the task of the test session in Experiment 3 was changed to an online test to exclude the possibility that the results of Experiments 1 and 2 were caused by how the tasks were specifically conducted: in the test session of Experiments 1 and 2, participants were asked to do a pen-and-paper test. In that task, participants were able to take their time thinking before making responses. To reduce the influence of this design on the results, participants in Experiment 3 were tested to segment words using an online test in the test session. By this design, participants did not have enough time to make strategic responses. We could exclude the possibility that participants in Experiments 1 and 2 made strategic responding based on their explicit recollection of having recently encountered the trained words. If the results of Experiments 1 and 2 were not caused by individual differences, and were not caused by task in the test session, Experiment 3 should replicate the major findings of Experiments 1 and 2.

Experiment 1

In Experiment 1, we investigated whether encountering words frequently in the auditory modality (listening) can affect word processing in the visual modality (reading). In the training session, Chinese participants listened to some words a few times in a short period of time. In the test session, we tested whether the frequently listened words are more likely to be segmented as words when segmenting an OAS, which has been shown to be sensitive to the word frequencies of the two words.

Method

Participants

Thirty college students (17 females and 13 males) participated in Experiment 1. Their ages ranged from 19 to 30 years (M = 23.4 years, SE = 0.09). All the participants were native Chinese speakers and had a normal or corrected-to-normal vision. The study was approved by the ethics committee of the XXX [name deleted to maintain the integrity of the review process]. The study conforms to the standards of Declaration of Helsinki. All participants gave their informed consent prior to their inclusion in the study.

Procedure

Each participant finished three sessions of the experiment: a training session and two test sessions. The first test session was conducted immediately after the training session, and the second test session was conducted about one week after the training session (M = 7.03 days, SE = 0.06, ranging between 6 and 8 days).

In the training session, participants did four blocks of experiments. In each block, they listened to all of the trained target words twice and did one of the following four tasks. In the first task, participants were asked to judge whether the two-character word they heard was a noun by pressing one key on the keyboard (the “1” key for nouns and the “0” key for other words). In the second task, they were asked to judge whether the two-character word they heard was a living thing. In the third task, they were asked to assess the familiarity of two-character words on a scale from 1 to 5. In the last task, they were asked to evaluate the semantic correlation between the two 2-character words on a scale of 1 to 5. For each trial of the experiment, participants listened to a word twice before doing the task. The order of the words was randomized within a block for each participant; in total, each target word was listened to eight times by each participant.

In each of the test sessions, participants were asked to do a pen-and-paper test. They were asked to draw a line at the boundaries of words in 40 OASs. To ensure that participants conducted the experiment seriously, four filler trials were included. In these fillers, none of the two characters made up a word (e.g., "实中左," in which no characters can be combined to form a word), so that participants would not mark any word boundary in the test.

Materials

We chose 40 two-character words as target words. These words were divided into two groups (e.g., "白板," means "white board"), with 20 words in each group. These words were low-frequency words (M = 0.37 occurrences per million, SE = 0.05, ranging between 0.04 and 0.89), and the properties were controlled between these two groups (see Table 1)[1]. These words were randomly assigned to two groups (groups A and B). Half of the participants experienced all words in group A but did

[1] The frequency of 33 of these target words was comparable (M=1.07 occurrences per million, SE=0.05, ranging between 0.03 and 12.10) in another word frequency corpus (Cai & Brysbaert, 2010). The other 7 targets were not listed in Cai and Brysbaert (2010).

not experience any words in group B in the training session. In contrast, the other half participants experienced all words in group B but did not experience any words in group A in the training session. With this design, any word was assigned to the training group for half participants and was assigned to the untrained group for the other half of the participants. By doing so, we can exclude the influences of word properties on the results to the largest extent.

In the test session, all of the 40 target words were embedded in OAS as the right-side word (e.g., "旁白板," in which the left-side word "旁白" means *voice over*, and the right-side word "白板" means *white board*). In half of these OASs, the right-side words were two-character words trained in the training session; in the other half, neither the left-side words nor the right-side words were trained in the training session. The word frequency of the left-side words (AB; $M = 0.69$ occurrences per million, $SE = 0.03$, ranging between 0.14 and 0.99) was significantly higher than that of the right-side words (BC; $M = 0.37$ occurrences per million, $SE = 0.04$, ranging between 0.04 and 0.89, $t(40) = 6.58$, $p < .001$). The left-side words and right-side words were both nouns. The number of strokes did not differ significantly between the first (A; $M = 8.65$, $SE = 0.35$) and the third characters (C; $M = 8.83$, $SE = 0.38$, $t(40) = 0.32$, $p = .749$). The first characters and the third characters can both be single-character words independently. Even though the words were infrequent, the characters of the words are known in their written form to the participants.

To examine how readers segmented the OASs without training, we presented them in isolation to another 18 participants (none of these participants participated in the main experiment). The results of this test showed that the probability of the middle word being segmented to the left-side word was higher than 0.6.

Results and Discussion

The results were analyzed using generalized linear mixed-effects models (GLMMs). The results of word segmentation (whether the first two characters constitute a word or the last two characters constituted a word) in the test sessions were treated as the independent variable; training type (trained or untrained), test session (immediate or 7-days interval), and their interaction were entered as fixed factors; and the participants and items were entered as crossed random effects, including intercepts and slopes (Baayen et al., 2008). Following Barr et al. (2013), we used the maximal model that could converge. We first constructed a model with a maximal random factor structure. When the maximum model did not converge, a zero-correlation parameter model was used, and we removed the random components that produce the minimum variance.

We used the *glmer* function from the lme4 package (Bates et al., 2018) in the *R* Environment for statistical computing (R Development Core Team, 2020) and reported regression coefficients (*b*s, which estimate the effect size), standard errors (*SE*s), *z*-values, and corresponding p-values (Table 2). The probability of A-BC segmentation was significantly higher in the trained condition (M = 0.47, SE = 0.03) than that in the untrained condition (M = 0.36, SE = 0.03, b = 0.33, SE = 0.07, z = 4.51, p < .001). The main effect of the test session was not significant (immediate condition: M = 0.42, SE = 0.03; 7-days interval condition: M = 0.40, SE = 0.03; b = -

0.03, *SE* = 0.06, *z* = -0.53, *p* = .595). Finally, the interaction between training and test session was significant (*b* = 0.11, *SE* = 0.05, *z* = 2.19, *p* = .029). We used the *emmeans* package (Lenth, 2022) in the *R* Environment for simple effect analysis, and the results showed that the training effect decreased with time. In the immediate test session, the probability of A-BC segmentation was larger in the trained condition (*M* = 0.50, *SE* = 0.04) than in the untrained condition (*M* = 0.33, *SE* = 0.03; *b* = -0.88, *SE* = 0.18, *z* = -4.93, *p* <.001). The difference between the trained condition and untrained condition was also significant in the 7-days interval test session (trained condition: *M* = 0.44, *SE* = 0.04; untrained condition: *M* = 0.36, *SE* = 0.03, *b* = -0.43, *SE* = 0.18, *z* = -2.40, *p* =.017), even though the effect size was smaller.

The results showed that recent word encountering experiences in the auditory modality impact word processing in the visual modality. The results of Experiment 1 showed that the results of OAS segmentation were affected by whether the right-side word was experienced in the training session. The OASs were more likely to be segmented as A-BC if BC was listened to frequently in the training session than otherwise. These results suggest that word encountering experience might be stored in some shared circuits for auditory modality and visual modality. We will discuss this issue in more detail in General Discussion.

Experiment 2

Experiment 1 suggested that word encountering experience can transfer from the auditory modality to the visual modality. As we were interested in whether the recent word encountering experience from listening was comparable with that acquired through reading, in Experiment 2, we investigated training effects within the visual modality. In the training session, Chinese participants saw some words a few times in a short period of time.

Method

Participants

Thirty college students (19 females and 11 males) participated in Experiment 2. Their ages ranged from 19 to 28 years (*M* = 22.4 years, *SE* = 0.08). All the participants were native Chinese speakers and had normal or corrected-to-normal vision. None of them took part in Experiment 1.

Materials

The materials were the same as those used in Experiment 1. Different from Experiment 1, the target words were shown as visual presentation eight times (twice in each of the four blocks) in the training session.

Procedure

The procedure was identical to Experiment 1, except that the stimuli were visually presented in the training session.

Results and Discussion

One participant was excluded because he/she did not show up in the 7-days interval test session. Trials in which participants did not make a response (either because participants judged that none of the characters constituted a word or that the three characters constituted a single word) were excluded from the analysis, resulting in the exclusion of 0.3% of the trials.

The method of data analysis was identical to that of Experiment 1. The results of the *GLMMs* are shown in Table 2. The probability of A-BC segmentation was significantly higher in the trained condition (M = 0.47, SE = 0.03) than in the untrained condition (M = 0.32, SE = 0.02, b = 0.38, SE = 0.06, z = 6.14, p < .001), and it was also significantly higher in the immediate test session (M = 0.45, SE = 0.03) than in the 7-days interval test session (M = 0.34, SE = 0.03, b = -0.31, SE = 0.10, z =-3.17, p = .002). Moreover, we found that the interaction between training and time was significant (b = 0.10, SE = 0.05, z = 2.09, p = .037). We used the *emmeans* package (Lenth, 2022) in the *R* Environment for simple effect analysis, which showed that the probability of A-BC segmentation was larger in the trained condition than in the untrained condition in both immediate test session (trained condition: M = 0.54, SE = 0.04; untrained condition: M = 0.36, SE = 0.03, b = -0.96, SE = 0.15, z = -6.19, p <.001) and 7-days interval test session (trained condition: M = 0.39, SE = 0.04; untrained condition: M = 0.29, SE = 0.03, b = -0.55, SE = 0.16, z = -3.50, p <.001).

The results of Experiment 2 are generally similar to those of Experiment 1. To directly compare the effects of word experience in different modalities, we conducted an analysis for the two experiments in a single model. In this model, training type, test session, modalities (trained in auditory modality or visual modality), and their interaction were entered as fixed effects, and the participants and items were entered as crossed random effects, including intercepts and slopes (Baayen et al., 2008). We report regression coefficients (*b*s, which estimate the effect size), standard errors (*SE*s), *z*-values, and corresponding p-values (Table 3).

The results showed a significant training effect and time effect, as analyzed separately in the two experiments before. The probability of A-BC segmentation was significantly higher in the trained condition than in the untrained condition ($b = 0.35$, $SE = 0.05$, $z = 6.63$, $p < .001$), and it was also significantly higher in the immediate test session than in the 7-days interval test session ($b = -0.16$, $SE = 0.06$, $z = -2.81$, $p = .005$). The main effect of modality was not significant ($b = 0.002$, $SE = 0.09$, $p = .984$), and modality only interacted with test session ($b = -0.14$, $SE = 0.06$, $p = .016$). Although the probability in the immediate test session was similar for two training types ($M = .42$ for listening training condition, $M = .45$ for the reading training condition), the probability of A-BC segmentation was higher in the listening training condition ($M = .40$) than in the reading training condition ($M = 0.34$) in the 7-days interval test session. The interaction between training type and training session was also significant ($b = 0.10$, $SE = 0.04$, $p = .003$). The training effect was slightly stronger in the immediate test session ($M = .17$) than in the 7-day test session ($M = .09$). None of the other interactions were significant.

Experiment 3

In the test session of Experiments 1 and 2, participants were asked to do a pen-and-paper test. In that task, participants were able to take their time thinking before making responses. Thus, it still left open the possibility of strategic responding based on their explicit recollection of having recently encountered the trained words.

In Experiment 3, participants were tested to segment words using an online test in the test session. By doing so, participants did not have enough time to make strategic responses. Moreover, we investigated training effects in different modalities using a within-participant design. In the training session, Chinese participants either saw or heard target words a few times in a short period of time.

Method

Participants

Thirty college students (23 females and 7 males) participated in Experiment 3. Their ages ranged from 19 to 35 years (M =24.6 years, SE = 0.12). All the participants were native Chinese speakers and had normal or corrected-to-normal vision. None of them took part in Experiments 1 or 2.

Materials

The materials were the same as those used in Experiment 1. Different from Experiments 1 and 2, each participant experienced three conditions in the training session. One-third of target words were shown in visual modality, one-third in auditory modality, and another one-third were not trained. From 40 targets in Experiments 1 and 2, we randomly removed one trial and divided the remaining 39 into three equal parts at random. Thus, there were 13 trials in each condition.

Procedure

The procedure was identical to Experiment 1, except that the stimuli were presented in three conditions (visual training, auditory training, and untrained) in the training session. As in Experiments 1 and 2, the target words were divided into three equal-numbered groups and were balanced across conditions and participants. With this design, any word was assigned to the auditory training group for one third of participants, was assigned to the visual training group for one third of participants, and was assigned to the untrained group for the other one third of participants.

In each of the test sessions, participants were asked to do an online word segmentation test as accurately and as quickly as possible. They saw a three-character string and were asked to press two buttons on a button box (Microsoft SideWinder Game Pad) to indicate whether the left two characters or the right two characters constituted a word. If none of the two characters constituted a word, they were instructed to press another button.

Results and Discussion

Trials in which participants did not make a response (either because participants judged that none of the characters constituted a word or that the three characters constituted a single word) were excluded from the analysis, resulting in the exclusion of 0.4% of the trials.

The results of word segmentation probability were analyzed using generalized linear mixed-effects models (*GLMMs*), and RTs were analyzed using linear mixed-effects models (*LMMs*)

Following the suggestions of Schad, Vasishth, Hohenstein, and Kliegl’s (2020), we performed the following planned contrasts: the effect of auditory training on the immediate test, the effect of auditory training on the 7-days test, the effect of visual training on the immediate test, the effect of visual training on the 7-days test, and the main effect of test session. The participants and items were entered as crossed random effects, including intercepts and slopes (Baayen et al., 2008). Following Barr et al. (2013), we used the maximal model that could converge. We first constructed a model with a maximal random factor structure. When the maximum model did not converge, a zero-correlation parameter model was used, and we removed the random components that produce the minimum variance. The segmentation and RTs results are shown in Table 4 and the results of the *GLMMs* and *LMMs* are shown in Tables 5 and 6.

RTs

The RTs were significantly longer in the immediate test session (M = 1,350 ms, SE = 12) than in the 7-days interval test session (M = 1,070 ms, SE = 7, b = 0.25, SE = 0.06, t = 4.37, p < .001). This difference in RTs might reflect the fact that participants have been more familiar with the task in the 7-days interval test session.

In the immediate test session, RTs were significantly shorter in the auditory training condition (M = 1,290 ms, SE = 5.9) than in the untrained condition (M = 1484 ms, SE = 21.8, b = -0.05, SE = 0.03, t = -2.05, p = .040), and it was also significantly shorter in the visual training condition (M = 1,279 ms, SE = 6.8) than in the untrained condition (M = 1,484 ms, SE = 21.8, b = -0.08, SE = 0.03, t = -2.28, p = .028). Shorter RTs for trained words reflect the influence of recent word encountering experiences on word processing, with the recently encountered word being processed faster. In the

7-days interval test session, no significant differences were found at the reaction time under the different conditions in the 7-days interval test session. The results of RTs showed that word encountering experience in listening affects the time that it needs for visual word processing. However, the training effect on RTs decayed fast so that we could only observe the effect in the immediate test session, but not in the 7-days interval test session.

The Probability of A-BC Segmentation

The probability of A-BC segmentation in the immediate test session (M = 0.55, SE = 0.005) was greater than in the 7-days interval test session (M = 0.51, SE = 0.005, b = 0.33, SE = 0.40, z = 0.20, p = .05).

In the immediate test session, the probability of A-BC segmentation was significantly higher in the auditory training condition (M =0.61, SE = 0.01) than in the untrained condition (M = 0.41, SE = 0.01, b = 1.02, SE = 0.20, z = 5.16, p < .001), and it was also significantly higher in the visual training condition (M = 0.65, SE = 0.01) than in the untrained condition (M = 0.41, SE = 0.01, b = 1.25, SE = 0.22, z = 5.74, p < .001).

In the 7-days interval test session, the probability of A-BC segmentation was significantly higher in the auditory trained condition (M =0.56, SE = 0.01) than in the untrained condition (M = 0.44, SE = 0.01, b = 0.59, SE = 0.16, z = 3.64, p < .001), and like in the immediate test session, it was also significantly higher in the visual trained condition (M = 0.53, SE = 0.01) than in the untrained condition (M = 0.44, SE = 0.01, b = 0.47, SE = 0.19, z =2.44, p = .014).

The results of Experiment 3 generally replicated the findings of Experiments 1 and 2. In Experiment 3, participants conducted on online word segmentation task. Using this task, participants need to respond as quickly as possible. In this situation, participants were not likely to make strategic responses. The replicated findings in Experiment 3 suggest that the impact of word encountering experience in auditory modality affect visual word processing is not solely caused by an untimed pen-and-paper test.

General Discussion

In the present study, we found that encountering a word frequently in a short period of time in the auditory modality can affect the segmentation of OASs in Chinese reading. In Experiment 1, participants listened to low-frequency target words frequently in the training session. These trained words were more easily segmented as words in the immediate test session. This effect was long-lasting and still observed after 7 days. In Experiment 2, participants read the same low-frequency target words in the training session, and a similar effect as in Experiment 1 was observed in the immediate test session and was still observed after 7 days. Because the segmentation of OASs is sensitive to word frequency, we argue that word encountering experience in the auditory modality can be transferred to the visual modality, so that it can affect the segmentation of OASs.

In Experiment 3, we used a within-subjects design, in which each participant received both auditory and visual training tasks. The possibility of participants responding strategically to the OAS segmentation was also further excluded from the online test task scores. Experiment 3 replicated the results of Experiments 1 and 2. Cross-modal effect of auditory training and the within-modality effect of visual training were significant in both the immediate test session and the 7-day interval test session.

These results have important implications for word processing models. First, the findings of the present study cannot be explained by those models that assume distinct mental lexicons (Morton, 1979, 1982). According to these models, listening and reading use different lexicons; therefore, listening to some words should only affect the identification threshold for that word in the listening lexicon. However, as OAS segmentation only involves reading and only uses the reading lexicon, the logogen models should not expect the transfer effect of word encountering experience. The results of the cross-modal transfer effect of word encountering experience suggest that word frequency information is represented at some location that is shared by both the auditory modality and the visual modality.

The results also have important implications for the triangle model. If word frequency information would only be represented as the weight of the links between phonology and semantics, the word frequency effect would not be expected to be transferred between modalities. According to the triangle model, readers encode semantic information through the direct semantic pathway and the phonologically mediated pathway. However, as Perfetti et al. (2005) argued, adult Chinese readers usually do not use the phonologically mediated pathway during Chinese reading. If word frequency information acquired from listening would only be represented as weights of the connection between phonology nodes and semantic nodes, it would not be expected to influence word processing during reading because adult Chinese readers do not use that pathway. Therefore, a word encountering experience from listening must be represented in some other place than in the connections between phonology nodes and semantic nodes.

Instead, word encountering experience might be stored in a place shared by listening and reading, such as in the semantic subnetwork. Harm and Seidenberg (2004) indicate connections between the semantic nodes and the cleanup nodes, and this subnetwork is used for both listening and reading. Therefore, it is likely that the word frequency information is stored as links between the semantic nodes and cleanup nodes. The argument that word frequency information is stored in a shared location by reading and listening is also supported by the finding that the training effects in the visual modality and the auditory modality were comparable. Otherwise, the training effect on visual word processing when the words were trained in the auditory modality should not be comparable with that when the words were trained in the visual modality.

Another possibility for cross-modal transfer effect of word encountering experience could be the result of coactivation of orthographic and phonological word forms during the training and/or test phase. Indeed, some studies have shown that the visual form is activated during spoken word perception in Chinese reading (Qu & Damian, 2017). The activated visual form might strengthen form-to-meaning mappings for both phonological and orthographic word representations, possibly leading to equivalent training effects in different training conditions.

In the present study, the cross-modal training effect was still observed 7 days after training. It is a noteworthy long-time continuation of the learning effect. In a study on long-time priming, researchers conducted a study on long-time priming by manipulating semantics, and the time interval of long-time priming reached an average of 20 minutes later (Rodd et al., 2013). The paradigm used in the present study may be more sensitive, so that we could observe the training effect even after 7 days.

Word frequency corpora are important to choose stimuli for experimental studies and also important to build computation models; however, previous corpora have usually only relied on written text. The findings of the present study suggest that the experience of spoken language should also be taken into account when constructing word frequency corpora. Indeed, some studies have used subtitle corpora as a proxy for spoken language (Cai & Brysbaert, 2010; New et al., 2007), and others have shown that word frequency information based on subtitle corpora explained more variances of lexical decision times than that only based on written materials (Brysbaert & New, 2009).

The present study has some limitations. First, we used the word segmentation task to investigate how word encountering experience in the auditory modality affects reading. Future work may need to use a natural reading task to investigate this question. Second, we only investigated how experience in the auditory modality affects reading. Future studies may consider training words in reading and test word segmentation in continuous speech.

Conclusion

The results of the present study showed that word encountering experiences in listening can be transferred to reading, affecting word segmentation during Chinese reading. Interestingly, the training effect in the auditory modality is comparable with the training effect in the visual modality. These results suggest that word encountering experiences should be stored at a location in the mind that is used for both listening and reading. Based on the results of the current study, this information is likely to be stored in the weights between the "cleanup" nodes and semantic nodes in the triangle model (Harm & Seidenberg, 2004).

Disclosure statement.
The authors report there are no competing interests to declare.

Data availability statement.

The data that support the findings of this study are openly available in OSF at https://osf.io/x9dc6/?view_only=3adbcf8e31dd4039a269f3046557acde

Reference

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*(3), 445-459. https://doi.org/10.3758/bf03193014

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology, 4,* 328.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). *Parsimonious Mixed Models* (arXiv:1506.04967). arXiv.

Bonin, P., & Fayol, M. (2002). Frequency effects in the written and spoken production of homophonic picture names. *European Journal of Cognitive Psychology*, *14*(3), 289-313. https://doi.org/10.1080/09541440143000078

Broadbent, D. E. (1967). Word-frequency effect and response bias. *Psychological Review*, *74*(1), 1-15. https://doi.org/10.1037/h0024206

Brown, G. D., & Watson, F. L. (1987). First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory & Cognition*, *15*(3), 208-216. https://doi.org/10.3758/bf03197718

Brysbaert, M. (2022). Word Recognition II: Phonological Coding in Reading. *The Science of Reading: A Handbook*, 79–101.

Brysbaert, M., & Ghyselinck, M. (2006). The effect of age of acquisition: Partly frequency related, partly frequency independent. *Visual Cognition*, *13*(7-8), 992-1011. https://doi.org/10.1080/13506280544000165

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*(4), 977-990. https://doi.org/10.3758/BRM.41.4.977

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). The impact of word prevalence on lexical decision times: Evidence from the Dutch Lexicon Project 2. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(3), 441-458. https://doi.org/10.1037/xhp0000159

Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica*, *104*(2), 215-226. https://doi.org/10.1016/s0001-6918(00)00021-4

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE, 5*(6), e10729. https://doi.org/10.1371/journal.pone.0010729

Chen, J., Lin, T.-J., Ku, Y.-M., Zhang, J., & O'Connell, A. (2018). Reader, Word, and Character Attributes Contributing to Chinese Children's Concept of Word. *Scientific Studies of Reading*, *22*(3), 209–224.

Cheng, C.-M. (1981). Perception of Chinese characters. *Acta Psychologica Taiwanica*.

Chua, F. K. (1999). Phonological recoding in Chinese logograph recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(4), 876.

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology*, *42*(4), 317-367.

Dehaene, S. (2009). *Reading in the brain: The science and evolution of a human invention*. Penguin Viking.

Daneman, M., & Stainton, M. (1991). Phonological recoding in silent reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(4), 618.

Donnenwerth-Nolan, S., Tanenhaus, M. K., & Seidenberg, M. S. (1981). Multiple code activation in word recognition: Evidence from rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, *7*(3), 170.

Drewnowski, A., & Healy, A. F. (1982). Phonetic factors in letter detection: A reevaluation. *Memory & Cognition*, *10*(2), 145–154. https://doi.org/10.3758/BF03209216

Ellis, A. W., Young, A. W., & Anderson, C. (1988). Modes of word recognition in the left and right cerebral hemispheres. *Brain and Language, 35*(2), 254-273.

Ferrand, L., Meot, A., Spinelli, E., New, B., Pallier, C., Bonin, P., Dufau, S., Mathot, S., & Grainger, J. (2018). MEGALEX: A megastudy of visual and auditory word recognition. *Behavior Research Methods, 50*(3), 1285-1307. https://doi.org/10.3758/s13428-017-0943-1

Forster, K. I., & Chambers, S. M. (1973). Lexical Access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12*(6), 627-635. https://doi.org/10.1016/s0022-5371(73)80042-8

Gardner, M. K., Rothkopf, E. Z., Lapan, R., & Lafferty, T. (1987). The word frequency effect in lexical decision: Finding a frequency-based component. *Memory & Cognition, 15*(1), 24-28. https://doi.org/10.3758/bf03197709

Gilbert, R. A., Davis, M. H., Gaskell, M. G., & Rodd, J. M. (2018). Listeners and readers generalize their experience with word meanings across modalities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(10), 1533–1561.

Grainger, J., & Ferrand, L. (1994). Phonology and orthography in visual word recognition: Effects of masked homophone primes. *Journal of Memory and Language, 33*(2), 218-233.

Grainger, J., Diependaele, K., Spinelli, E., Ferrand, L., & Farioli, F. (2003). Masked repetition and phonological priming within and across modalities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(6), 1256–1269. https://doi.org/10.1037/0278-7393.29.6.1256

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111*(3), 662.

Harris, M., & Coltheart, M. (1986). *Language processing in children and adults: An introduction.* Taylor & Francis.

Hoosain, R. (1991). Cerebral Lateralization of Bilingual Functions After Handedness Switch in Childhood. *The Journal of Genetic Psychology*, *152*(2), 263–268.

Huang, L., & Li, X. (2020). Early, but not overwhelming: The effect of prior context on segmenting overlapping ambiguous strings when reading Chinese. *Quarterly Journal of Experimental Psychology. 73*(9), 1382-1395.https://doi.org/10.1177/1747021820926012.

Huang, L., Staub, A., & Li, X. (2021). Prior context influences lexical competition when segmenting Chinese overlapping ambiguous strings. *Journal of Memory and Language*, *118*, 104218. https://doi.org/10.1016/j.jml.2021.104218

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word-processing during eye fixations in rading: Effects of word frequency. *Perception & Psychophysics, 40*(6), 431-439. https://doi.org/10.3758/bf03208203

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329-354. https://doi.org/10.1037/0033-295x.87.4.329

Kempley, S. T., & Morton, J. (1982). The effects of priming with regularly and irregularly related words in auditory word recognition. *British Journal of Psychology, 73*(4), 441-454.

Kligler, N., & Gabay, Y. (2023). A cross-modal investigation of statistical learning in developmental dyslexia. *Scientific Studies of Reading.* Advance online publication. https://doi.org/10.1080/10888438.2023.2166413

Lenth, R.V. (2022). emmeans: Estimated marginal means, aka Least-Squares means. R package version 1.7.2. https://CRAN.R-project.org/package=emmeans

Li, X., Huang, L., Yao, P. & Hyönä, J. (2022). Universal and specific reading mechanisms across different writing systems. *Nature Reviews Psychology*.10.1038/s44159-022-00022-6.

Li, X., Rayner, K., & Cave, K. R. (2009). On the segmentation of Chinese words during reading. *Cognitive Psychology*, *58*(4), 525–552.

Li, X., Zhao, W., & Pollatsek, A. (2012). Dividing lines at the word boundary position helps reading in Chinese. *Psychonomic Bulletin & Review, 19*(5), 929–934.

Liu, P., Li, W., Lin, N., & Li, X. (2013). Do Chinese Readers Follow the National Standard Rules for Word Segmentation during Reading? *PLoS ONE*, *8*(2), e55440.

Liu, X., Vermeylen, L., Wisniewski, D., & Brysbaert, M. (2020). The contribution of phonological information to visual word recognition: Evidence from Chinese phonetic radicals. *Cortex*, *133,* 48–64.

Ma, G., Li, X., & Rayner, K. (2014). Word segmentation of overlapping ambiguous strings during Chinese reading. *Journal of Experimental Psychology: Human Perception and Performance, 40*(3), 1046–1059. https://doi.org/10.1037/a0035389

McPherson, W. B., Ackerman, P. T., & Dykman, R. A. (1997). Auditory and Visual Rhyme Judgements Reveal Differences and Similarities Between Normal and Disabled Adolescent Readers. *Dyslexia*, *3*(2), 63–77. https://doi.org/10.1002/(SICI)1099-0909(199706)3:2<63::AID-DYS49>3.0.CO;2-Q

Mok, L. W. (2009). Word-superiority effect as a function of semantic transparency of Chinese bimorphemic compound words. *Language and Cognitive Processes*, *24*(7–8), 1039–1081.

Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 116–133.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review 76*(2): 165-178.

Morton, J. (1979). Facilitation in word recognition: Experiments causing change in the logogen model. In P. Kolers (Ed.), *Processing of visible language* (pp. 259–268). Springer.

Morton, J. (1982). Disintegrating the lexicon: An information processing approach. *Perspectives on Mental Representation*, 89-109.

Muneaux, M., & Ziegler, J. (2004). Locus of orthographic effects in spoken word recognition: Novel insights from the neighbour generation task. *Language and Cognitive Processes*, *19*(5), 641–660. https://doi.org/10.1080/01690960444000052

New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics, 28*(4), 661-677. https://doi.org/10.1017/s014271640707035x

Pastizzo, M. J., & Carbone, R. F. (2007). Spoken word frequency counts based on 1.6 million words in American English. *Behavior Research Methods*, *39*(4), 1025-1028. https://doi.org/10.3758/bf03193000

Pattamadilok, C., Knierim, I. N., Duncan, K. J. K., & Devlin, J. T. (2010). How does learning to read affect speech perception? *Journal of Neuroscience*, *30*(25), 8435–8444.

Pattamadilok, C., Knierim, I., & Devlin, J. T. (2008). Orthographic influence on speech processing: A combined TMS and behavioral study. *Brain Stimulation*, *1*(3), 321–322.

Pattamadilok, C., Perre, L., Dufau, S., & Ziegler, J. C. (2009). On-line Orthographic Influences on Spoken Language in a Semantic Task. *Journal of Cognitive Neuroscience*, *21*(1), 169–179.

Patterson, K., Shewell, C., Coltheart, M., Sartori, G., & Job, R. (1987). Speak and spell: Dissociations and word-class effects. M. Coltheart, G. Sartori, & R. Job

(Eds.), *The Cognitive Neuropsychology of Language (*pp. 273-294). Lawrence Erlbaum Associates, Inc.

Perfetti, C. A., Liu, Y., & Tan, L. (2005). The lexical constituency model: Some implications of research on Chinese for general theories of reading. *Psychological Review*, *112*(1), 43-59. https://doi.org/10.1037/0033-295X.112.1.43

Pollatsek, A., Drieghe, D., Stockall, L., & de Almeida, R. G. (2010). The interpretation of ambiguous trimorphemic words in sentence context. *Psychonomic Bulletin & Review*, *17*(1), 88–94.

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 148–162. https://doi.org/10.1037/0096-1523.18.1.148

Qu, Q., & Damian, M. F. (2017). Orthographic effects in spoken word recognition: Evidence from Chinese. *Psychonomic Bulletin & Review*, *24*(3), 901–906.

R Development Core Team. (2020). *R: A language and environment for statistical computing* (Version 4.0.0) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org

Rack, J. P. (1985). Orthographic and phonetic coding in developmental dyslexia. *British Journal of Psychology*, *76*(3), 325–340.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372-422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, *14*(3), 191-201. https://doi.org/10.3758/bf03197692

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, *81*(2), 275.

Rodd, J. M., Cai, Z. G., Betts, H. N., Hanby, B., Hutchinson, C., & Adler, A. (2016). The impact of recent and long-term experience on access to word meanings: Evidence from large-scale internet-based experiments. *Journal of Memory and Language, 87*, 16–37.

Rodd, J. M., Cutrin, B. L., Kirsch, H., Millar, A., & Davis, M. H. (2013). Long-term priming of the meanings of ambiguous words. *Journal of Memory and Language*, *68*(2), 180–198.

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 9*(5), 487-494.

Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(6), 546.

Seidenberg, M.S., & McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523-568.

Shen W., & Li X. (2012). 中文阅读中词优效应的特异性. *Chinese Science Bulletin, 57*(35), 3414–3420.

Tan, L. H., & Perfetti, C. A. (1997). Visual Chinese Character Recognition: Does Phonological Information Mediate Access to Meaning? *Journal of Memory and Language*, *37*(1), 41–57.

Tyler, M. D., & Burnham, D. K. (2006). Orthographic Influences on Phoneme Deletion Response Times. *Quarterly Journal of Experimental Psychology, 59*(11), 2010–2031. https://doi.org/10.1080/17470210500521828

Van Assche, E., Duyck, W., & Gollan, T. H. (2016). Linking recognition and production: Cross-modal transfer effects between picture naming and lexical decision during first and second language processing in bilinguals. *Journal of Memory and Language, 89*, 37-54. http://dx.doi.org/10.1016/j.jml.2016.02.003

Van Orden G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition, 15*(3), 181–198. https://doi.org/10.3758/bf03197716

Ventura, P., Morais, J., Pattamadilok, C., & Kolinsky, R. (2004). The locus of the orthographic consistency effect in auditory word recognition. *Language and Cognitive Processes, 19*(1), 57–95. https://doi.org/10.1080/01690960344000134

Wei, W., Li, X., & Pollatsek, A. (2013). Word properties of a fixated region affect outgoing saccade length in Chinese reading. *Vision Research*, *80*, 1-6. https://doi.org/10.1016/j.visres.2012.11.015

Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology*, *1*(1), 59–85.

Xu, Y., Pollatsek, A., & Potter, M. C. (1999). The activation of phonology during silent Chinese word reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(4), 838.

Yan, G., Tian, H., Bai, X., & Rayner, K. (2006). The effect of word and character frequency on the eye movements of Chinese readers. *British Journal of Psychology*, *97*, 259-268. https://doi.org/10.1348/000712605x70066

Yao, P., Alkhammash, R., & Li, X. (2021). Plausibility and Syntactic Reanalysis in Processing Novel Noun-noun Combinations During Chinese Reading: Evidence From Native and Non-native Speakers. *Scientific Studies of Reading*, 1–19. https://doi.org/10.1080/10888438.2021.2020796.

Zecker, S. G. (1991). The orthographic code: Developmental trends in reading-disabled and normally-achieving children. *Annals of Dyslexia*, *41*(1), 178–192. https://doi.org/10.1007/BF02648085

Zecker, S. G., Tanenhaus, M. K., Alderman, L., & Siqueland, L. (1986). Lateralization of lexical codes in auditory word recognition. *Brain and Language*, *29*(2), 372–389. https://doi.org/10.1016/0093-934X(86)90055-6

Zhao, J., Li, T., Elliott, M. A., & Rueckl, J. G. (2018). Statistical and Cooperative Learning in Reading: An Artificial Orthography Learning Study. *Scientific Studies of Reading, 22*(3), 191–208.

Zhou, X., & Marslen-Wilson, W. (1999). Phonology, Orthography, and Semantic Activation in Reading Chinese. *Journal of Memory and Language, 41*(4), 579–606.

Zhou, X., Marslen-Wilson, W., Taft, M., & Shu, H. (1999). Morphology, Orthography, and Phonology Reading Chinese Compound Words. *Language and Cognitive Processes, 14*(5–6), 525–565.

Table 1

The Properties of Stimuli in Experiments 1 and 2

	Group A		Group B		*t**	*p*
Properties	*M*	*SD*	*M*	*SD*		
Frequency of AB	.67	.03	.71	.06	-0.45	.652
Frequency of BC	.39	.05	.34	.05	0.55	.583
Strokes of A	8.8	5.33	8.5	4.47	0.43	.671
Strokes of C	8.45	9.21	9.2	5.01	-0.89	.379

Note. The data were obtained from the lexicon of common words in contemporary Chinese (Lexicon of Common Words in Contemporary Chinese Research Team, 2008). The frequency measure stands for frequency per million words

* The degree of freedom of t test was 38.

Table 2

The Probability of A-BC Segmentation in Experiments 1 and 2

	Immediate test session		7-days interval test session	
	Trained	Untrained	Trained	Untrained
Experiment 1	.50 (.04)	.33 (.03)	.44 (.04)	.36 (.03)
Experiment 2	.54 (.04)	.36 (.03)	.39 (.04)	.29 (.03)

Note. SEs are presented in parentheses.

Table 3

Results of the Generalized Linear Mixed-effects Models in Experiments 1 and 2

Effects	*b*	*SE*	*z*	*p*
Experiment 1				
Trained	**0.33**	**0.07**	**4.51**	**<.001**
Test session	-0.03	0.06	-0.53	.595
Interaction	**0.11**	**0.05**	**2.19**	**.029**
Experiment 2				
Trained	**0.38**	**0.06**	**6.14**	**<.001**
Test session	**-0.31**	**0.10**	**-3.17**	**.002**
Interaction	**0.10**	**0.05**	**2.09**	**.037**
Cross-Experiments				
Trained	**0.35**	**0.05**	**6.63**	**<.001**
Test session	**-0.16**	**0.06**	**-2.81**	**.005**
Modality	0.001	0.09	0.02	.984
Trained - Test session Interaction	**0.10**	**0.04**	**2.96**	**.003**
Trained - Modality Interaction	0.03	0.04	0.57	.571
Test session - Modality Interaction	**-0.14**	**0.06**	**-2.41**	**.016**
Three-way Interaction	-0.004	0.04	-0.10	.919

Note. Significant effects are indicated in bold.

Table 4

Results of Experiments 3

	Immediate test session			7-days interval test session		
	Auditory Trained	Visual Trained	Untrained	Auditory Trained	Visual Trained	Untrained
Probability of A-BC Segmentation	.61 (.01)	.65 (.01)	.41 (.01)	.56 (.01)	.53 (.01)	.44 (.01)
RTs	1290 (5.9)	1279 (6.8)	1484 (21.8)	1055 (4.6)	1057 (5.3)	1073 (7.7)

Note. SEs are presented in parentheses.

Table 5

Statistics Results of Probability of A-BC Segmentation of Experiments 3

Effects		*b*	*SE*	*z*	*p*
Auditory Trained	Immediate test session	**1.02**	**0.20**	**5.16**	**<.001**
	7-days interval test session	**0.59**	**0.16**	**3.64**	**<.001**
Visual Trained	Immediate test session	**1.25**	**0.22**	**5.74**	**<.001**
	7-days interval test session	**0.47**	**0.19**	**2.44**	**0.015**
Test Session		0.40	0.20	1.94	.051

Note. Significant effects are indicated in bold.

Table 6

Statistics Results of RTs of Experiments 3

Effects		*b*	*SE*	*z*	*p*
Auditory Trained	Immediate test session	**-0.05**	**0.03**	**-2.05**	**.040**
	7-days interval test session	-0.01	0.03	-0.48	.630
Visual Trained	Immediate test session	**-0.78**	**0.03**	**-2.28**	**0.028**
	7-days interval test session	-0.01	0.03	-0.42	0.671
Test Session		**0.25**	**0.06**	**4.37**	**<.001**

Note. Significant effects are indicated in bold.

Appendix

Table A.1

Materials of Experiment 1

Target words	OAS
白板	旁**白板**
彩印	迷**彩印**
笔试	粉**笔试**
表针	课**表针**
网眼	球**网眼**
肉松	肥**肉松**
虫牙	甲**虫牙**
针眼	钢**针眼**
迷彩	沉**迷彩**
春药	立**春药**
节操	字**节操**
风骨	古**风骨**
根须	病**根须**
盒饭	烟**盒饭**
花卷	麻**花卷**
底牌	锅**底牌**
票根	验**票根**
矿灯	盐**矿灯**
轮班	油**轮班**
雷暴	春**雷暴**
雨刷	阵**雨刷**
路牌	套**路牌**
校服	驾**校服**
木鱼	积**木鱼**
色狼	粉**色狼**

钱庄	零**钱庄**

Target words	OAS
牌局	冒**牌局**
堂姐	讲**堂姐**
村姑	渔**村姑**
印堂	彩**印堂**
台步	吧**台步**
败将	破**败将**
草包	香**草包**
单亲	货**单亲**
休书	午**休书**
宫刑	龙**宫刑**
凶案	帮**凶案**
束腰	花**束腰**
骨盆	排**骨盆**
油彩	豆**油彩**

Note: The target words in the overlapping ambiguous strings are marked in bold.