Peer reviewed version

Link to published version (if available):
[10.1037/xlm0000520](10.1037/xlm0000520)

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research
General rights

Orthographic effects in second-language spoken word recognition

Qingqing Qu[1,2], Zhanling Cui[3], Markus F. Damian[4]

[1]Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

[2]Department of Psychology, University of Chinese Academy of Sciences, Beijing, China

[3] Department of Education, Hebei Normal University, Shijiazhuang, Hebei, China

[4]School of Experimental Psychology, University of Bristol, United Kingdom

Word count: 5,316 (excl. abstract and references)

Address for correspondence:
Qingqing Qu
Key Laboratory of Behavioral Science
Institute of Psychology
Chinese Academy of Sciences, Beijing, China
16 Lincui Road, Chaoyang District, Beijing, China
100101
China
Tel: +86-10-64888629
Fax: +86-10-64872010
Email: quqq@psych.ac.cn

Abstract

Substantial evidence from alphabetic languages suggests a role of orthography in the processing of spoken words. Recently, orthographic effects have also been reported for speakers of Mandarin Chinese, a language with a non-alphabetic orthographic system. However, most existing work has focused on native spoken word processing. For native languages, the acquisition of orthographic information occurs much later than that of spoken language. Contrary to native languages, non-native languages are typically learned in a form where the acquisition of orthographic information parallels or even precedes the one of phonological information. Given these differences, it is unclear whether orthographic codes are similarly accessed from spoken words in non-native languages. To explore this issue, Tibetan-Chinese bilinguals judged whether or not Chinese spoken words presented in pairs were related in meaning. Some of the unrelated word pairs were orthographically related, and critically, this orthographic overlap induced a significant increase in response latencies. Compared to previous results from L1 listeners with the identical procedure, the orthographic effect for L2 listeners was more pronounced. These findings indicate that orthographic information is involuntarily accessed in native and non-native spoken-word recognition alike, and that it may play a more important role in the latter compared to the former."

Key words: Orthography; Nonnative spoken word recognition; Semantic judgment task; Chinese

Orthographic effects in second-language spoken word recognition

Most researchers concur with the view that phonological information constrains orthographic input processing in visual word recognition (e.g., Frost, 1998; Spinks, Liu, Perfetti, & Tan, 2000; Zhou & Marslen-Wilson, 1999). The complementary claim, i.e. that orthographic information is involved in spoken word recognition, is perhaps less intuitive but it is supported by a growing number of studies (e.g., Seidenberg & Tanenhaus, 1979; Ziegler & Ferrand, 1998). This observation has motivated so-called "on-line" accounts according to which bidirectional links between orthography and phonology allow for the involuntary cross-activation of orthographic codes whenever phonological representations are accessed (e.g,. Pattamadilok, Perre, Dufau, & Ziegler, 2009; Slowiaczek, Soltano, Wieting, & Bishop, 2003). Alternatively, according to "offline-restructuring" accounts (Perrea, Pattamadilokc, Montanta, & Zieglera, 2009), the acquisition of literacy leads to the restructuring of an individual's phonological representations, with orthographic effects in speech resulting from access to "phonographic" representations. The phonological restructuring view dovetails with the fact that developmentally, phonological knowledge is acquired much earlier than orthographic knowledge. Different from native languages, however, most bilingual individuals learned their non-native (L2) representations based on both written and spoken codes. Because for the non-native language of bilinguals, the acquisition of orthographic information parallels or even precedes the one of phonological information, orthographic and phonological codes might be more intimately interlinked than is the case for native languages. Alternatively, orthographic L2 representations might be <u>less</u> stable in bilinguals than they are for L1 speakers because relative exposure is reduced in L2 compared to L1, and therefore orthographic effects in speech processing might be less pronounced. The aim of the present study was to investigate 1) whether orthography is activated in spoken word recognition of non-native listeners, and 2) if so, whether orthographic effects differ between native and non-native listeners.

Over the last few decades, studies on the processing of spoken language have demonstrated orthographic effects in a variety of tasks. In a seminal study, Seidenberg and Tanenhaus (1979) showed that in an auditory rhyme judgment task, judgments were made faster for word pairs that were spelled similarly (e.g., tie/pie) than for word pairs spelled differently (e.g., tie/rye). Frauenfelder, Segui and Dijkstra (1990) demonstrated that in phoneme detection performed on spoken French words, detection of /k/ took longer than of /p/, supposedly because /k/ in French has more possible spellings ("c", "cc", "k", "ck", "qu", etc.) than /p/ (see Dijkstra, Roelofs, & Fieuws, 1995 for a similar finding in Dutch). The validity of such meta-phonological tasks has been questioned on the basis that orthographic effects might reflect sophisticated response strategies, rather than automatic activation of orthographic information (e.g., Cutler, Treiman, & van Ooijen, 1998; Damian & Bowers, 2000). However, orthographic effects have also been demonstrated in priming paradigms conducted in the auditory modality which are less open to this criticism. In these tasks, it is tested whether an orthographic relation between a prime and a target modulates priming (Chéreau, Gaskell, & Dumay, 2007; Slowiaczek, Soltano, Wieting, & Bishop, 2003; Perre, Midgley, & Ziegler, 2009). For example, in an auditory lexical decision task, Jakimik, Cole and Rudnicky (1985) found significant priming only when primes and targets shared both orthographic and phonological information, but not when they overlapped only phonologically, or only orthographically. Further evidence for orthographic effects comes from studies in which it is orthographic properties of the target word which are manipulated. In an auditory lexical decision task, Ziegler and Ferrand (1998) found that the orthographic consistency of English words affected response latencies: responses were faster for consistent words (with a rhyme that can be spelled in only one way) than for inconsistent words (with a rhyme that can be spelled in multiple ways). Such orthographic consistency effects have now been documented in various languages (English: Miller & Swick, 2003; Portuguese: Ventura, Morais, Pattamadilok, & Kolinsky, 2004; French: Pattamadilok, Morais, Ventura, &

Kolinsky, 2007), and not only in unimpaired readers, but also in alexic patients (e.g., Miller & Swick, 2003).

Apart from meta-phonological or lexical decision tasks, orthographic effects also emerge in semantic tasks. For instance, Pattamadilok, Perre, Dufau, and Ziegler (2009) combined electroencephalography (EEG) measurements with a task in which participants were instructed to press a response button when a given word belonged to a semantic category, and to withhold their response otherwise. On no-go trials, either orthographically consistent or inconsistent words were presented. EEG results revealed a clear orthographic consistency effect, the onset of which was time-locked to the position of the inconsistency in the spoken word. The effect occurred at a relatively early point in time, suggesting that orthography affects lexical access, rather than a later decisional or postlexical stage. Overall, a large number of studies adopting a range of experimental manipulations and paradigms support the claim that orthographic information influences spoken word recognition (see e.g. Hallé, Chéreau, & Segui, 2000; Taft, Castles, Davis, Lazendic, & Nguyen-Hoan, 2008; Ziegler & Ferrand, 1998; Ziegler, Ferrand, & Montant, 2004 for additional evidence; however, see Cutler & Davis, 2012, for the view that such effects could reflect strategic adaptations to a particular task environment, and Mitterer & Reinisch, 2015, arguing that orthographic effects might not extend to conversational speech).

The reviewed evidence stems from speakers of languages which employ alphabetic orthographic systems. In these languages, the mapping between spelling and sound is systematic (even in languages such as English with a high degree of irregularity), and so cross-talk between orthographic and phonological representations is perhaps unsurprising. An interesting question that arises concerns languages with non-alphabetic scripts, in which sublexical spelling-sound correspondences are largely absent. For instance, in Chinese, the basic spoken unit is the syllable which maps onto a written character. To exemplify, 汉语

(Chinese, /han4yu3/) is a disyllabic word that is composed of two characters (汉, /han4/, Chinese & 语, /yu3/, language), which in turn, are composed of one or more radicals (e.g., 语 consists of two radicals "讠" &"吾"), which, in turn, are composed of strokes. Due to the lack of sublexical links between sound and spelling, words can be orthographically similar (for instance, two words might share an orthographic radical) while being phonologically unrelated. Given the relative lack of the sublexical correspondences between sound and spelling which characterize non-alphabetic languages, one might therefore speculate that compared to alphabetic languages, processing links between phonology and orthography should be correspondingly reduced.

Does spoken language processing in non-alphabetic languages still involve orthographic co-activation? Zou, Desroches, Liu et al. (2012) manipulated orthographic overlap between prime-target pairs in an auditory lexical decision task, and found that orthographic similarity modulated ERP amplitudes, with significantly reduced N400 amplitudes when a target was preceded by an orthographically similar prime. Chen, Chao, Chang et al. (2016) investigated effects of homophone density (referred to the number of characters sharing the same pronunciation, thus taken to index orthographic variation at the character level) and orthographic consistency (referred to the consistency from phonology to orthography at the level of radicals, thus assumed to reflect orthographic variation at the radical level,) on Chinese spoken word recognition via EEG. Participants were asked to judge whether or not a spoken word represented an animal. In EEG, orthographic consistency, modulated the amplitude of N400, whereas homophone density modulated a late positive component (LPC). Finally, we (Qu & Damian, 2016) reported results from a semantic relatedness judgment task in which native Mandarin speakers judged whether or not spoken word pairs were related in meaning. Word pairs were either semantically related (e.g., 错误, /cuo4wu4/, error-准确, /zhun3que4/, exact), orthographically related (i.e., shared a radical, e.g., 错误, /cuo4wu4/, error-蜡烛, /la4zhu2/, candle), or

unrelated. Results showed that relatedness judgments were made faster for word pairs that were semantically related than for unrelated word pairs. Critically, on semantically unrelated pairs, response latencies were slower for word pairs which were orthographically related (i.e., shared a radical) than for unrelated pairs. Overall, the evidence suggests that orthographic activation in spoken word processing is ubiquitous across languages with alphabetic and non-alphabetic scripts.

All studies reviewed so far explored speech processing in the participants' native language. A related – and largely unexplored – issue is whether orthographic effects also arise when non-native listeners process speech. Native (L1) speakers of a given language acquired their representations of spoken language at a very early age in life, whereas orthographic codes were added only much later. By contrast, most non-native (L2) speakers learned their L2 representations based on both written and spoken codes. Because the acquisition of orthographic information parallels or even precedes the one of phonological information, lexical codes are likely to be co-structured with input from the two subsystems (Veivo & Järvikivi, 2013). Based on this argument, one might predict that orthographic effects in non-native language processing would be more pronounced than those documented in native listeners (see evidence reviewed above). On the other hand, it is possible that orthographic effects in non-native speech processing might be less pronounced than in native language processing because orthographic representations in L2 are not as solid as those in L1 (see De Bot & Lowie, 2010; Jiang, 2000).

So far, only very few studies have investigated the role of orthography in non-native spoken word recognition. Veivo and Järvikivi (2013) adopted a masked cross-modal priming paradigm in which Finnish learners of French carried out lexical decisions on spoken French words which were preceded by briefly presented and masked visual primes. The relationship between primes and target words was manipulated

such that primes were (i) repetitions of target words (e.g., a spoken target "stage" preceded by visual prime

"stage"; primes were phonologically and orthographically identical to target words); (ii) non-word

pseudohomophones of target words (e.g., "staje"; primes were phonologically identical to targets, but only

partially orthographically related); or (iii) non-word controls which were unrelated to targets. Hence,

repetition and pseudohomophone primes were equated in terms of phonological relatedness with the

target, and the reasoning was that if priming effects relative to the baseline condition are larger for

repetition than for pseudohomophone primes, this should reflect a role of orthography in L2 spoken

processing. Results were consistent with this prediction: repetition primes generated larger priming effects

than pseudohomophone primes. Further, this benefit depended on participants' L2 proficiency, with larger

effects for more proficient participants than for less proficient ones. Converging results were reported by

Veivo, Järvikivi, Porretta and Hyönä (2016) who used the "visual world" task in which Finnish learners of

French, as well as French native speakers, were asked to match spoken words with printed words on a

computer screen while their eye movements were recorded. Fixations on French target words were

compared to those on competitors for which either orthographic overlap was held constant while

phonological overlap was manipulated (Experiment 1; "base" vs "bague" or "bain"), or phonological overlap

was held constant while orthographic overlap was manipulated (Experiment 2; "mince", /mɛs/ vs. "mite" ,

/mit/ or "mythe" , /mit/). Clear orthographic effects did not emerge either in the L1 nor the L2 listener group.

However, an effect of proficiency was observed for orthographic overlap over time within each trial, with an

orthographic effect emerging for higher but not for lower L2 proficiency listeners.

These and related findings (e.g., Mishra & Singh, 2014) provide initial evidence that even in L2 speech

processing, orthographic codes are co-activated. However, a possible objection might be that in tasks of this

type, primes or competitors are visually presented, which generates an orthographic context that caused

participants to strategically activate information about the spelling of the spoken stimuli. Hence, results

might not be representative of the way in which listeners process speech in the real world. A further

potential problem specific to the use of the auditory lexical decision task used in Veivo and Järvikivi (2013) is

that participants might strategically generate an orthographic image of the spoken word to facilitate

decisions on the lexical status of the word (e.g., Cutler, Treiman, & Van Ooijen, 1998; Pattamadilok, Perre,

Dufau, & Ziegler, 2009), which if true, would again imply that results are not representative of speech

processing outside the lab.

The present study was designed to address these issues. We used the same semantic judgement task

featured in Qu and Damian (2016): participants were presented with two successive spoken Mandarin words

("prime" and "target"), and they were asked to judge via a key press response whether or not the two words

were semantically related. Pairs were either semantically related, orthographically related, or unrelated,

hence expected responses were "yes" for semantically related pairs, and "no" for orthographically related

and unrelated pairs. However, whereas Qu and Damian (2016) used native Chinese Mandarin speakers, here

we tested a group of Tibetan non-native Chinese speakers. Based on our previous findings with L1 listeners,

we expected that responses would be faster for semantically related word pairs than unrelated ones,

because negative responses generally involve more complex decision than "yes" responses (e.g., Gomez,

Ratcliff, & Perea, 2007; Wu & Thierry, 2010). Critically, finding that responses on semantically unrelated pairs

are modulated by orthographic relatedness would provide evidence for an influence of orthographic

knowledge on non-native spoken word recognition, and a comparison of this effect with the one reported in

Qu and Damian for L1 listeners would provide further evidence about whether the effect is more or less

pronounced, or similar, in L1 and L2 listeners. Hence, our task 1) was exclusively based on spoken codes and

avoided explicit use of orthography altogether; 2) arguably, is more natural than lexical decisions or

metalinguistic analysis (semantic judgments are generally assumed to be strategy-free because they require

listeners to retrieve the meaning of spoken words but deflect from explicit analysis of form representations;

e.g., Pattamadilok et al., 2009, for a review); 3) involved a target language in which spelling and sound can be

largely dissociated (i.e., a non-alphabetic orthographic system).

Method

Participants

Thirty-two students from Affiliated National College of Hebei Normal University participated in the

study (17 females). All participants reported normal hearing, normal or corrected-to-normal vision and no

history of neurological or language problems. All were unbalanced late bilinguals, with Tibetan language[1] as

their L1 and Mandarin Chinese as L2. All of the participants had begun learning Chinese at school at a mean

age of 5.8, and claimed to be medium to high proficient in Chinese. See Table 1 for additional background

information.

Materials and Design

The materials and design were identical to Qu and Damian (2016). Stimuli consisted of 105 disyllabic

word quartets, each involving one target, and three prime words. In each set, the target was paired with a

prime word which was either (1) semantically related (but unrelated in word form; e.g., 枕头, pillow,

/zhen3tou2/ - 被子, quilt, /bei4zi/), (2) orthographically related (but phonologically and semantically

unrelated; prime and target shared one radical with the target, e.g., 破裂, break, /po4lie4/ - 被子, quilt,

/bei4zi/), or (3) unrelated in phonology, orthography or meaning (e.g., 酸奶, yogurt, /suan1nai3/ - 被子,

---

[1] Tibetan is a language spoken primarily on the high plateau north of the Himalayas. The Tibetan writing system is derived from an Indian prototype. Tibetan is only very distantly related to Chinese in terms of its spoken and written properties.

quilt, /bei4zi/). Form overlap in participants' L1 (i.e., in the Tibetan language) was avoided in all

combinations. Prime words were matched on the frequency of the first character, word frequency, stroke

numbers of the first character, and stroke numbers of the whole word ($Fs < 1$, $ps > .299$); the mean frequency

of target words, semantically related prime words, orthographically related prime words, and unrelated

prime words was 19.24, 20.97, 21.35, 13.89 per million respectively as determined by the Chinese Linguistic

Data Consortium (2003) norms. Moreover, semantic relatedness between orthographically related word pairs

and unrelated word pairs was matched.[2] The full set of stimuli is available online at

http://eyemind.psych.ac.cn/enpublication.html.

From these materials three lists were created such that each participant was presented with 35 trials in

each of the three conditions, and no word was repeated in either of the lists. In this way, across the three

lists each target word was presented with the three different prime types. The three types of prime-target

combinations were distributed across three blocks. Each participant was presented with three blocks of 35

trials within each block, for a total of 105 trials.

Stimuli were recorded by a female native speaker of Chinese, at a sampling rate of 44 kHz. The mean

duration of primes was 812 ms ($SD = 78$) for semantic primes, 814 ms ($SD = 88$) for orthographic primes, and

804 ms ($SD = 76$) for unrelated primes. There were no significant differences in duration across the three

---

[2] We collected semantic rating scores on a seven-point Likert scale for all pairs of words from a group of 57 Tibetan-Chinese participants (1 = "not related at all," and 7 = "closely related"). The semantic rating scores were 4.24, 1.64, and 1.60 for semantically related, orthographically related, and unrelated word pairs respectively. The difference in semantic relatedness between semantically related word pairs and the other conditions was significant, $ps < .001$, but critically, there was no difference between the orthographically related and the unrelated condition ($p > .284$). Hence, the stimuli were semantically well matched across orthographically related and unrelated word pairs, which had also been the case in ratings collected from native Chinese speakers and reported in Qu et al. (2016).

.

types of prime words (ps > .363). The mean duration of targets was 810 ms (SD = 74).

Procedure

The procedure was identical to Qu and Damian (2016). Stimuli were presented using E-Prime 1.1 software (Psychology Software Tools, Pittsburgh, PA). Participants were first instructed about the task, and subsequently were presented with six practice trials (3 semantically related trials, 3 unrelated trials). After the practice, three experimental blocks of 35 trials each were presented. Stimuli were presented in random sequence through headphones. There were short breaks between blocks, and the next block started after participants indicated that they were ready to continue. On each trial, participants were presented with a sequence consisting of a fixation cross (500 ms), a spoken prime word, ISI (1000 ms) and a spoken target word. The intertrial interval was 1,000 ms. Participants were asked to decide as quickly as possible whether or not the two words they heard were semantically related by pressing the key "f" if they were semantically related, and "j" otherwise. Response latencies were measured from the onset of the target word to the participants' response. Each experimental task session lasted approximately 20 minutes. At the end of each experiment, the experimenter asked the participants whether they had noticed any associations between the two words they heard, other than the obvious semantic relatedness on some trials. None of the participants reported having noticed orthographic overlap between words.

Results

Trials with incorrect responses (10.5%) and trials with responses faster than 200 ms or slower than 3,000 ms (2.6%) were excluded from the response time analysis. As shown in Table 2, relative to the unrelated condition, response latencies exhibited a substantial facilitatory effect (96 ms) of semantic relatedness, and a substantial inhibitory effect (75 ms) of orthographic relatedness.

A linear mixed-effects model analysis (Baayen, Davidson, & Bates, 2008; Bates, 2005) was conducted on the response latencies, with the fixed factor prime type, and by-participant and by-item random intercepts and slopes for prime type. Model fitting was carried out by initially specifying a model that only included by-participant and by-item random intercepts and slopes for prime type, and then we incrementally added the fixed factor prime type into the model. The model comparison, expressed as a chi-squared statistic, showed that the addition of prime type improved the initial model fit, $\chi2 = 23.65$, $p < .001$. Analysis of the model revealed that compared to the unrelated condition, latencies in the semantically related condition were significantly faster, $b = -102.76$, $SE = 39.86$, $t = -2.58$, $p = .014$, whereas in the orthographically related condition they were slower, $b = 87.65$, $SE = 22.42$, $t = -3.91$, $p < .001$. A parallel analysis was conducted on the errors but with a binomial family due to the binary nature of the data (Jaeger, 2008). Here, the "maximal" model with by-participant and by-item random intercepts and slopes failed to converge, so a simpler model with random intercepts only was used. The linear mixed-effects model analysis showed that the addition of prime type significantly improved the initial model fit, $\chi^2 = 22.89$, $p < .001$, and compared to the unrelated condition, the semantically related condition showed significantly higher error rates, $b = 0.69$, $SE = 0.15$, $z = 4.75$, $p < .001$, as did the orthographically related condition, $b = 0.47$, $SE = 0.15$, $z = 3.12$, $p = .002$.

The results reported by Qu and Damian (2016) with the identical procedure but conducted on L1 listeners are also included for comparison in Table 2. Numerically, the orthographic effect is almost twice as large in L2 than in L1, whereas the semantic effect appears to be quite similar in size. To explore this pattern, we performed a joint analysis on the results from both experiments. Data were analysed for orthographic and semantic effects separately. First, we focused on orthographic effects. Model fitting was carried out by initially specifying a model that included group (L2 vs L1) and prime type (orthographically related vs. unrelated) were included as fixed factors, and random by-participant and by-items intercept and slope adjustments were specified

for prime type (but not for the between-participants factor group; see Barr et al., 2013), and then the group x prime type interaction was added into the model. The addition of the interaction significantly improved the initial model fit, $\chi2 = 5.47$, $p = .019$. Analysis of the most complex model indicated that the effect of prime type, $b = -44.60$, $SE = 17.59$, $t = -2.54$, $p = .012$, group, $b = 381.98$, $SE = 54.18$, $t = 7.05$, $p < .001$, and critically group x prime type interaction, $b = -43.84$, $SE = 18.57$, $t = -2.36$, $p = .018$, were significant. Hence, the hypothesis that the orthographic effect was more pronounced for L2 than for L1 listeners received statistical support.[3] Parallel analyses conducted on the semantic effects showed that the addition of the group x prime type interaction did not improve the model fit, $\chi2 = 0.02$, $p = .881$. The main effect of prime type, $b = 97.46$, $SE = 32.81$, $t = 2.97$, $p = .004$, and group, $b = 332.75$, $SE = 37.72$, $t = 8.82$, $p < .001$, were significant, whereas the interaction between both factors was not, $b = 6.36$, $SE = 42.85$, $t = 0.15$, $p = .882$. The lack of the interaction between semantic relatedness and group indicates that the magnitude of semantic effects was similar between the two groups.

A parallel joint analysis was conducted on the errors. The error analysis conducted on orthographic effects showed that the group x prime type interaction did not improve the model fit, $\chi2 = 1.07$, $p = .300$; and the main effect of group was significant, $b = 0.60$, $SE = 0.28$, $z = 2.13$, $p = .033$, whereas the main effect of prime type and the interaction between both factors was not, $|z|<1.12$, ps > .289. The error analysis conducted on semantic effects showed that the group x prime type interaction did not improve the model fit, $\chi2 < 1$, $p = .489$; and the main effect of prime type, $b = -0.50$, $SE = 0.19$, $z = -2.65$, $p = .008$, and group, $b = 0.65$, $SE = 0.17$, $z = 3.91$, $p < .001$, were significant, whereas the interaction

---

[3] Note that overall responses for L2 listeners were 349 ms slower (~25%) than that for L1 listeners. It could be argued that orthographic effects in the L2 group were larger than in the L1 group simply because the former had overall slower response latencies than the latter. To explore this possibilty, we computed, for L1 and L2 participants separately, the correlation between by-participant overall latency, and the size of the orthographic effect, but found none ($r = -.03$, $p = .88$, and $r = +.01$, $p = .96$, respectively). This makes it unlikely that group differences in overall latencies, rather than the L1/L2 distinction, underlies the difference in the size of the orthographic effect.

between both factors was not, $|z|$<1.00, ps > .488.

Discussion

A growing number of studies provide support for the view that native language spoken word processing involves the activation of orthographic information. By contrast, very few studies have investigated the role of orthography in non-native spoken word recognition, and virtually all such studies were conducted with speakers of alphabetic languages in which the close correspondence between spelling and sound makes it difficult to tease the two apart. Furthermore, as noted in the Introduction, existing studies presented stimuli in visual and auditory modalities, and a possible objection to the inferences drawn from these studies is that doing so creates an artificial context which biases participants towards activating orthographic information of spoken words which might not be representative of speech processing under more naturalistic circumstances. In the present study, we adopted an exclusively auditory task in which Tibetan-Chinese bilinguals were presented with two spoken Chinese words and judged whether the two words were related in meaning or not. On unrelated pairs, a subset were orthographically related, and as in a previous study which targeted native Chinese listeners (Qu & Damian, 2016), for L2 listeners in the current study orthographic overlap modulated response latencies. The finding constitutes clear evidence that orthography plays a role in second-language spoken comprehension. This inference converges with the one drawn by recent studies conducted with alphabetic languages (e.g., Veivo & Järvikivi, 2013) but it extends the effect to non-alphabetic languages. The activation of orthography appeared to be unconscious and automatic because when questioned at debriefing, none of the participants reported being aware of the hidden orthographic overlap.

In the present study, semantic judgement responses were faster for semantically related than for

unrelated word pairs. This pattern probably reflects the fact that positive responses are generally made

faster than negative ones. By contrast, orthographic overlap on semantically unrelated trials induced a

significant increase in reaction times, compared to the baseline condition. This inhibitory effect might arise

because relatedness between the automatically activated orthographic codes creates a response conflict for

the semantic relatedness decision (i.e., orthographic overlap suggests a positive response, whereas the

required response is negative), thus resulting in longer response latencies. The general pattern parallels

results of previous studies. For instance, Thierry and Wu (2004) showed that when Chinese-English bilinguals

performed semantic relatedness judgments on visually presented English word pairs, overlap in Chinese

orthographic properties slowed down responses and elicited greater error rates.

In conjunction with the earlier findings from Qu and Damian (2016), the present results suggest that

spoken word processing is modulated by orthographic overlap in both L1 and L2 listeners. In processing

terms, how exactly such effects come about remains at present speculative. In the current task, individuals

carry out semantic relatedness judgements on spoken word pairs, and critical (orthographically unrelated or

related word pairs) require a negative response, i.e., the detection that the word pair is semantically

unrelated. Presumably, individuals temporarily hold the two spoken words in short-term memory, scan for

semantic relatedness, perhaps via priming resulting from conceptual overlap, and carry out a positive

response if such priming is detected, and a negative response otherwise. For instance, semantic priming

between the two word pairs could result in elevated activation levels of the corresponding word nodes, and

individuals might base their response on monitoring these activation levels. In some current processing

models of visual and spoken word recognition such as the Bimodal Interactive Activation Model (BIAM;

Grainger & Ferrand, 1994), phonological and orthographic pathways are closely and bidirectionally

connected. If so, then in our task, orthographic overlap for a semantically unrelated word pair might induce

priming between the two lexical nodes which "mimics" semantic priming by raising activation levels of the lexical nodes, therefore making it more difficult to carry out the required negative response then for an entirely unrelated word pair, resulting in slower latencies.

It is noteworthy that the magnitude of the orthographic effect was larger (almost twice as large) in L2 than in L1. We speculate that this pattern arises from the difference in the strength of connections between orthography and phonology in native and nonnative language processing. Speakers acquire the phonological forms of their native language first, and map phonology onto orthographic forms much later on; by contrast, late bilinguals typically learn their L2 in instructional settings in which orthographic and phonological forms are acquired in conjunction. The early and significant exposure to orthography during the acquisition of second language could lead to stronger connection between orthography and phonology in L2 lexical knowledge, which explains why orthography evidently plays a more important role in second-language than in native-language spoken comprehension. Moreover, semantic effects were nearly identical in magnitude across both groups. This renders it unlikely that the differential orthographic effects arose from other potential confounding factors (such as overall response speed; see Footnote 3), which if true should have affected the magnitude of the semantic effects as well.

As outlined in the Introduction, it has recently been suggested that proficiency in L2 might affect the extent to which orthography affects speech processing in L2 (Veivo & Järvikivi, 2013; Veivo et al., 2016). To explore this possibility for our participants, we performed correlational analyses between participants' self-rated L2 proficiency levels, length of residence in a Chinese speaking area, or the age of acquisition for L2 (see Table 1) and the size of the orthographic effect, but found no significant correlations. Another way to evaluate the influence of proficiency in L2 on orthographic effects is to include the three proficiency factors in the linear

mixed-effects models.[4] The results showed that none of the three factors was significant, $|t| < 1.28$, $ps > .210$.

These null findings could be because the range of proficiency in our population was relatively restricted, or

alternatively because our measures of proficiency were not precise enough to capture an underlying relation.

Further research is needed to explore the possibility that the involvement of orthography in L2 speech

processing is affected by L2 proficiency.

Future research on this issue should additionally examine a potential role of homophone density in

Chinese spoken processing. Homophones are prevalent in Chinese; according to Yip (2000), only 23 out of

1273 heterotonic syllables (considering tones) have no homophonic characters. This might be the reason why

the modern Chinese lexicon gradually transformed from classical monosyllabic dominance into contemporary

disyllabic dominance. Although adopting words with two character could somewhat reduce the homophone

difficulty, the problem remains. Disyllabic homophones are not uncommon in modern Mandarin Chinese

(e.g., 公事, public affairs, /gong1shi4/-攻势, offensive, /gong1shi/-公式, formula, /gong1shi4/). The

homophone ambiguity is only resolved when the orthography of the character is determined, which might

render the involvement of orthography more pronounced in Chinese than in alphabetic languages. An

interesting direction for further research would be to examine whether orthographic effects are a function of

the homophone density of characters (i.e., the number of characters sharing the same sound). If homophone

ambiguity is relevant, one would predict more pronounced orthographic effects for spoken words written

with characters with high, than with low, homophone density: when a character has more homophones, its

orthography should play a greater role in determining its meaning than when a character has few

homophones.

---

[4] We would like to thank an anonymous reviewer for suggesting this additional analysis.

To conclude, the results of the present study indicate that nonnative listeners with a formal instruction background in L2 activate orthographic information during nonnative spoken word recognition, perhaps to a larger extent than L1 listeners do.

Acknowledgements

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory & Language, 59,* 390-412.

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

Bates, D. M., Maechler, M., & Dai, B. (2009). lme4: Linear mixed-effect models using S4 classes. R package version 0.999375-28. [Software]. Vienna, Austria: R Foundation for Statistical Computing.

Chéreau, C., Gaskell, M. G., & Dumay, N. (2007). Reading spoken words: orthographic effects in auditory priming. *Cognition, 102,* 341–360.

Chinese Linguistic Data Consortium. (2003). 现代汉语通用词表 [Chinese lexicon] (CLDC-LAC-2003-001). Beijing, China: Tsing-hua University, State Key Laboratory of Intelligent Technology and Systems, and Chinese Academy of Sciences, Institute of Automation

Cutler, A,. & Davis, C. (2012). An orthographic effect in phoneme processing, and its limitations. *Frontiers in Psychology, 3*:18.

Cutler, A., Treiman, R., & van Ooijen, B. (1998). *Orthografik inkoncistensy ephekts in foneme detektion?* In Proceedings of the Fifth international conference on Spoken Language Processing (pp. 2783_2786). Sydney: ICLSP.

Damian, M. & Bowers, J. (2010). Orthographic effects in rhyme monitoring tasks: Are they automatic? *European Journal of Cognitive Psychology, 22,* 106-116.

De Bot, K., & Lowie. W. (2010). On the stability of representations in the multilingual lexicon. In M. Pütz & L. Sicola (eds.) *Cognitive Processing in Second Language Acquisition*, pp. 117–134. Amsterdam; Philadelphia: John Benjamins.

Dijkstra, T., Roelofs, A., & Fieuws, S. (1995). Orthographic effects on phoneme monitoring. *Canadian Journal of Experimental Psychology, 49,* 264–271.

Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information-processing rate and amount: implications for group differences in response latency. *Psychological Bulletin, 125*, 777–799.

Frauenfelder, U. H., Segui, J., & Dijkstra, T. (1990). Lexical effects in phonemic processing: facilitatory or inhibitory? *Journal of Experimental Psychology: Human Perception & Performance, 16,* 77-91.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123,* 71–99.

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

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go lexical decision task. *Journal of Experimental Psychology: General, 136,* 347–369.

Hallé, P. A., Chéreau, C., and Segui, J. (2000). Where is the/b/in "absurde" [apsyrd]? It is in French listeners' minds. *Journal of Memory and Language, 43,* 618–639.

Jakimik, J., Cole, R. A., & Rudnicky, A. I. (1985). Sound and spelling in spoken word recognition. *Journal of Memory and Language, 24,* 165–178.

Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics, 21,* 47–77.

Miller, K. M., & Swick, D. (2003). Orthography influences the perception of speech in alexic patients. *Journal of Cognitive Neuroscience, 15,* 981-990.

Mishra, R. K., & Singh, N. (2014). Language non-selective activation of orthography during spoken word

    processing in Hindi–English sequential bilinguals: an eye tracking visual world study. *Reading and*

    Writing, 27, 129-151.

Mitterer, H., & Reinisch, E. (2015). Letters don't matter: no effect of orthography on the perception of

    conversational speech. *Journal of Memory and Languages, 85,* 116-134.

Pattamadilok, C., Morais, J., Ventura, P., & Kolinsky, R. (2007). The locus of the orthographic consistency

    effect in auditory word recognition: further evidence from French. *Language and Cognitive Processes,*

    22, 1-27.

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,* 169–179.

Perre, L., Midgley, K., & Ziegler, J. C. (2009). When beef primes reef more than leaf: Orthographic information

    affects phonological priming in spoken word recognition, *Psychophysiology, 46,* 739–746.

Perre, L., Pattamadilok, C., Montant, M., & Ziegler, J. C. (2009). Orthographic effects in spoken language: On-

    line activation or phonological restructuring? *Brain Research, 1275,* 73–80.

Qu, Q., & Damian, F. M. (2016). Orthographic effects in spoken word recognition: Evidence from Chinese.

    Psychonomic Bulletin & Review, DOI 10.3758/s13423-016-1164-9

Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of*

    Experimental Psychology: Human Learning and Memory, 5, 546–554.

Slowiaczek, L. M., Soltano, E. G., Wieting, S. J., & Bishop, K. L. (2003). An investigation of phonology and

    orthography in spoken-word recognition. *The Quarterly Journal of Experimental Psychology, 56*(2), 233–

    62.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults

and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 461–479.

Spinks, J. a, Liu, Y., Perfetti, C. a, & Tan, L. H. (2000). Reading Chinese characters for meaning: the role of phonological information. *Cognition, 76*(1), B1–B11.

Taft, M., Castles, A., Davis, C., Lazendic, G., & Nguyen-hoan, M. (2008). Automatic activation of orthography in spoken word recognition: pseudohomograph priming. *Memory and Language, 58,* 366–379.

Thierry, G., & Wu, Y. J. (2004). Electrophysiological evidence for language interference in late bilinguals. *Neuroreport, 15,* 1555–1558.

Veivo, O., & Järvikivi, J. (2013). Proficiency modulates early orthographic and phonological processingin L2 spoken word recognition. *Bilingualism: Language and Cognition, 16,* 864–883.

Veivo, O., Järvikivi, J., Porretta, V., & Hyönä, J. (2016). Orthographic activation in L2 spoken word recognition depends on proficiency: evidence from eye-tracking. *Frontiers in Psychology, 7*.

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,* 57-95.

Wu, Y. J., & Thierry, G. (2010). Chinese-English bilinguals reading English hear Chinese. The *Journal of Neuroscience, 30*, 7646–7651.

Yip, P.-C. (2000). *The Chinese lexicon: A comprehensive survey.* London: Routledge.

Zhou, X., & Marslen-Wilson, W. (1999). Phonology, orthography, and lexical semantic activation in reading Chinese. *Journal of Memory and Language, 41*, 579-606.

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review, 5,* 683–689.

Ziegler, J. C., Ferrand, L., & Montant, M. (2004). Visual phonology: the effects of orthographic consistency on

different auditory word recognition tasks. *Memory & Cognition, 32*, 732–741.

Table 1

Participants background information

Background information	Mean	Range
Age (in years)	18.3	17-20
Age of acquisition for L2 Chinese (in years)	5.8	2-9
L2 proficiency	5.0	3-6
Length of residence in a Chinese speaking area (in years)	5.0	3-14
Education level (in years)	12.0	12

Note: Participants represent a narrow range of proficiency levels, based on self-evaluations of their own proficiency levels (7 point scale, with 1 representing "not familiar at all with Chinese", and 7 representing "extremely familiar with Chinese")

Table 2

Response latencies (in milliseconds), error percentages, and their standard deviations in parentheses, by condition. Results from Qu and Damian (2016) are shown for comparison on the right hand side

Condition	Example prime-target pairs Target: 被子, "quilt", /bei4zi/	Response latencies	RT Effect	Errors	Errors Effect	Response latencies (errors) from Qu & Damian (2016)	RT Effect	Errors	Errors Effect
Unrelated	Prime: 酸奶, yogurt, /suan1nai3/	1443 (257)		7.3 (7.4)		1122 (144)		4.7 (3.8)	
Orthographically related	Prime: 破裂, break, /po4lie4/	1518 (259)	-75	11.0 (11.6)	-3.7	1162 (143)	-40	5.7 (5.0)	-1.0
Semantically related	Prime: 枕头, pillow, /zhen3tou2/	1347 (176)	96	13.2 (6.7)	-5.9	1022 (129)	100	7.4 (4.6)	-2.7

Note: Conditional means deviate very slightly from those reported in the original study because Qu and Damian reported means of means, whereas here we report means of raw latencies.