LANGUAGE LEARNING
A Journal of Research in Language Studies

Language Learning ISSN 0023-8333

EMPIRICAL STUDY

The Processing of English Derived Words by Chinese-English Bilinguals

Junmin Li,[a] Marcus Taft,[b] and Joe Xu[b]

[a]Zhejiang University City College and [b]University of New South Wales

This study examined the sensitivity of Chinese-English bilinguals to derivational word structure in English. In the first experiment, English monolinguals showed masked priming effects for prime-target pairs related both transparently (e.g., *hunter-HUNT*) and opaquely (e.g., *corner-CORN*) but not for those related purely in terms of form (e.g., *freeze-FREE*), whereas bilinguals showed priming in all three conditions. Furthermore, stronger form priming was found for bilinguals who were less experienced in English. A second experiment showed that bilingual participants found it harder to identify items as nonwords when the words possessed a suffix (e.g., *animalful*) than when they did not (e.g., *animalfil*), and this was true in terms of accuracy even for bilinguals with less exposure to English. Overall, these findings suggest that Chinese-English bilinguals, regardless of proficiency, have some sensitivity to morphological structure and that greater proficiency leads to priming effects that tend to pattern more like those of monolinguals.

Keywords bilingual lexical processing; Chinese-English bilinguals; derived words; morphological awareness; morphological processing

Introduction

It has been well established since the early work of Taft and Forster (1975) that derivational word recognition involves a morphological decomposition process (e.g., Clahsen & Neubauer, 2010; Clahsen, Felser, Neubauer, Sato, & Silva, 2010; Diependaele, Duñabeitia, Morris, & Keuleers, 2011; Silva & Clahsen,

---

The research reported in this article was conducted while Junmin Li was a visiting scholar at the University of New South Wales, funded by the China Scholarship Council in 2014 (File No. 201406320257). We thank Professor Li Degao for helping collect the data at Zhejiang University and acknowledge all the valuable suggestions made by the reviewers and editors.

Correspondence concerning this article should be addressed to Marcus Taft, School of Psychology, University of New South Wales, Sydney, Australia, NSW 2052. E-mail: m.taft@unsw.edu.au

Language Learning 00:0, xxxx 2017, pp. 1–27

DOI: 10.1111/lang.12247

2008; Ullman, 2004, 2005). This process has been revealed in lexical decision experiments using a masked priming paradigm (e.g., Longtin, Segui, & Hallé, 2003; Rastle, Davis, Marslen-Wilson, & Tyler, 2000; Rastle, Davis, & New, 2004) where recognition of a target word (e.g., *HUNT* or *CORN*), typically distinguished from prime words by using capital letters, is facilitated both when the prime is a complex word composed of that target plus a suffix (e.g., *hunter*) and when the prime is a pseudo-derived or opaque word (e.g., *corner*), which has the appearance of an affixed version of the target but is semantically unrelated to it. In contrast, a prime that begins with a word but does not end in a suffix-like string (e.g., *freeze*, where *ze* is not a suffix) does not facilitate the recognition of the target (i.e., *FREE*). These results have provided support for morpho-orthographic processing, in which decomposition of the letter string occurs simply because it appears to be morphologically complex. Studies using masked priming (e.g., Feldman et al., 2009) have also indicated that transparent prime–target pairs (e.g., *hunter-HUNT*) produced more priming than opaque ones (e.g., *corner-CORN*), suggesting an early involvement of a morphosemantic stage of processing. Even in these studies, though, opaque priming was nevertheless significant and typically greater than pure orthographic priming (e.g., *freeze-FREE*). So it is apparent that words like *corner* that only appear to have morphological structure are decomposed during their recognition.

Although this may be true for native English speakers, is it the case that second language (L2) learners rely on a similar mechanism to process complex words in their L2? In fact, there is empirical evidence for the idea that L2 learners place greater emphasis on whole-word retrieval in visual recognition of derived words than do native speakers (Clahsen & Neubauer, 2010; Clahsen et al., 2010; Silva & Clahsen, 2008; Ullman, 2004, 2005). For example, Silva and Clahsen (2008) examined in a series of masked priming experiments the performance of groups of bilingual speakers whose L2 was English. While native English speakers demonstrated a morphological priming effect that was of the same magnitude as identity priming (e.g., *neatness* primed *NEAT* to the same degree that *neat* primed *NEAT*, relative to the unrelated pair *dark-NEAT*), the L2 groups showed only partial priming for the morphologically related pairs, with a magnitude of priming that was intermediate between the unrelated and identity conditions. This outcome was observed for first language (L1) speakers of both German and Japanese, despite the differences in their orthographic similarity to English and even though both languages make use of affixation. The finding that masked morphological priming was smaller for L2 participants than for native speakers suggested that, compared with L1 processing, processing of L2 relies more on lexical storage than morphological decomposition.

Recently Heyer and Clahsen (2015) directly compared the degree of masked priming for native English speakers and late L2 learners at an advanced level of English for morphologically related (e.g., *darkness-DARK*) and orthographically related (e.g., *example-EXAM*) pairs. While native readers showed morphological priming and no form-based priming, L2 learners showed priming in both conditions. Heyer and Clahsen argued that early in the word-recognition process, L2 learners focus more on surface form than on morphological structure.

In contrast, however, other studies have suggested that adult L2 learners process derivationally complex words in the same way as native speakers (Diependaele et al., 2011; Portin, Lehtonen, & Laine, 2007; Zhang, Liang, Yao, Hu, & Chen, 2016). For example, Zhang et al. (2016) investigated morphosemantic activation in early visual recognition by comparing semantically transparent derived word pairs (*reporter-REPORT*), semantically opaque pairs (*corner-CORN*), and semantically related pairs (*choice-SELECTION*). Chinese learners of English showed a priming effect in the transparent condition when the stimulus onset asynchrony was 40 milliseconds, fully masking the prime from awareness, and a priming effect in the opaque condition only when the stimulus onset asynchrony was 80 milliseconds. In neither of these situations were priming effects found for semantically related words. It was concluded that Chinese L2 learners can employ rule-based decomposition based on both morphosemantic and morpho-orthographic information.

Diependaele et al. (2011) found a large masked priming effect in the transparent condition (e.g., *hunter-HUNT*), weaker opaque priming (e.g., *corner-CORN*), and no significant form priming (e.g., *turnip-TURN*) for native English speakers, with Spanish-English and Dutch-English bilinguals of varying levels of English proficiency showing the same graded pattern of facilitation. The form-priming effect was actually significant for the bilingual speakers, but it was weaker than the transparent priming effect. Similarly, a masked priming study by Duñabeitia, Dimitropoulou, Morris, and Diependaele (2013) found a pattern of priming with transparent words that was the same for monolinguals and bilinguals (Spanish-English and Basque-Spanish), regardless of how balanced the latter were in terms of proficiency in their L1 and L2. Such studies suggest that L2 learners, regardless of proficiency, automatically decompose complex words, just as native speakers do.

It might be considered surprising that L2 proficiency appeared to have no impact on processing. However, the low-proficiency groups of the Diependaele et al. (2011) and Duñabeitia et al. (2013) studies might have been sufficiently experienced in their L2 to have developed advanced processing strategies. In the Diependaele et al. study, the Spanish-English participants were considered more proficient in English than the Dutch-English participants, having higher scores on a reading test and an earlier age of acquisition. However, speaking ability did not differ between the two groups, which raised the possibility that even the Dutch participants were advanced learners. In the Duñabeitia et al. study, a language proficiency questionnaire showed that the Spanish-Basque bilinguals were highly and equally proficient in both languages while the Spanish-English bilinguals had a higher level of proficiency in Spanish than in English. However, that did not mean that the latter had a low level of English proficiency. In both these studies, then, it is possible that the less-proficient bilinguals were in fact quite proficient in their L2. In addition, proficiency was determined across different L1 backgrounds, potentially confounding proficiency with other factors.

There are other studies, however, that, despite examining L2 proficiency within a single population of bilinguals, have also found support for a decompositional strategy regardless of proficiency. Casalis, Commissaire, and Duncan (2015) employed a lexical decision task with a focus on pseudoword classification responses. Two groups of French-English bilinguals who had acquired English at the same age but who differed in their proficiency were tested. It was found that both groups were more likely to wrongly accept pseudowords made up of an embedded word and a suffix as real words (e.g., *proudy*) compared to other pseudowords, indicating that even lower-proficiency L1 speakers can be sensitive to morphological structure in that language.

The Current Study

This study examined Chinese-English bilinguals with different levels of proficiency to test whether they were able to decompose English derived words. In Chinese, over 75% of words are composed of more than one character (Kuo & Anderson, 2006), but very few of these characters function as an affix (Packard, 2000). Previous studies have shown that the recognition of inflected words is influenced by the typological distance between L1 and L2 (e.g., Basnight-Brown, Chen, Hua, Kostić, & Feldman, 2007; Portin et al., 2008; Vainio, Pajunen, & Hyönä, 2014). For example, Basnight-Brown et al. (2007) found that Chinese-English bilinguals showed less facilitation in irregular inflected word priming (e.g., *drawn-draw* or *swung-swing*) than did bilinguals with an Indo-European L1. An effect of typological distance may also be true when it comes to derived words, given that derivational suffixes are no more a feature of Chinese than are inflectional suffixes.

The impact of proficiency on morphological processing with Chinese-English bilinguals was examined by Liang and Chen (2014) who found significant behavioral priming of regular inflected words (e.g., *walked-WALK*) and evidence for morphological priming from highly proficient Chinese-English bilinguals in the form of an attenuated N400 component but not from less proficient L2 learners. Such research might not be relevant to the current study, however, because it looked at inflected rather than derived words. The linguistic difference between derivational and inflectional morphemes lies in the fact that the former create new words in the mental lexicon while the latter convey functional or grammatical influences on the stem. Thus, the processing of derived words and inflected words may well be different (e.g., see Taft, 2004). For example, sensitivity to inflectional information (as shown by Liang and Chen, 2014) might arise from learning a set of rules (e.g., "add an *s* to pluralize" in English) or from memorizing an inflectional paradigm (Silva & Clahsen, 2008), whereas sensitivity to derivational structures requires knowledge of relevant affixes and the contexts in which they are appropriately used. Given how little we know about the recognition of English derived words by Chinese-English bilinguals, the issue merits investigation.

Experiment 1

In the first experiment, a masked priming lexical decision task was adopted to investigate whether Chinese-English bilinguals of different proficiency levels are able to decompose English derived words regardless of semantic transparency in an early stage of visual word recognition. This paradigm was designed to reveal whether well-known words are automatically decomposed into morphemes. Not only were transparent prime–target pairs (e.g., *hunter-HUNT*) tested, but opaque ones were as well (e.g., *corner-CORN*). In addition, form-related pairs of words (e.g., *freeze-FREE*) were included to establish whether any priming effects observed were based purely on overlapping orthographic form or on morpho-orthographic decomposition. Both the native speakers and the bilinguals undertook the same masked priming task with the same set of materials.

The view that L2 learners rely on whole-word storage more than L1 speakers do (Clahsen & Neubauer, 2010; Clahsen et al., 2010; Silva & Clahsen, 2008; Ullman, 2004, 2005) predicts that L2 learners will have reduced or no priming effects compared to L1 speakers in all conditions. In contrast, the findings of Diependaele et al. (2011) and Duñabeitia et al. (2013) have suggested that L2 learners will show a similar pattern of responding to that of L1 speakers, namely, stronger priming effects for both transparent and opaque derived words than for

Table 1 Bilingual participants' background variables and vocabulary knowledge in Experiment 1

	Group C	Group I
Variable	*M* (*SD*)	*M* (*SD*)
Age	19.3 (0.9)	22.5 (0.9)
Age of acquisition	9.3 (2.1)	9.6 (2.0)
Years of learning English	10.2 (2.2)	11.8 (1.9)
Weekly use of Chinese (%)	92.9 (7.1)	60.2 (10.8)
Weekly use of English (%)	7.1 (5.1)	39.8 (20.9)
Vocabulary test score	10.8 (3.6)	14.6 (4.4)

form-related words. A third possibility also exits: Bilinguals might only behave like L1 speakers when they have higher L2 proficiency, with lower proficiency bilinguals showing no or reduced priming.

Method

Participants

English monolinguals[1] and Mandarin-English bilinguals took part in the experiment. The English monolingual group consisted of 43 undergraduate students at the University of New South Wales (UNSW) who received course credit for their participation. The mean age of monolingual participants was 19.82 years ($SD = 4.26$). In order to optimize the range of English proficiency among the bilinguals, two participant pools were drawn upon. The first consisted of 41 international students (majoring in engineering) with a Chinese background (Group I) studying at UNSW, where an overall International English Language Testing System score of 6.5 and above is required for admission. This group was therefore assumed to have a high level of English proficiency. They received A$10 for their participation. The second pool of bilingual participants consisted of 60 undergraduate students majoring in engineering at Zhejiang University (ZJU) in China (Group C), who received ¥20 for their participation. Given their background, this group was assumed to have a lower level of proficiency than Group I because the latter comprised mostly postgraduate students who would have spent more years studying English and had far greater exposure to an English-language environment. This assumption was confirmed by a questionnaire given to the bilingual participants to extract estimates of language use (see Table 1), where it was found that the two groups differed markedly in their estimated weekly use of English, $t(99) = 15.89$, $p < .001$,

$d = 3.19$. Table 1 also provides details about the age of the participants, as well as the age at which they first began to acquire English.

All bilingual participants also took a pen-and-paper English vocabulary test with 30 items arranged in order of increasing difficulty (Andrews & Hersch, 2010; see Appendix S1 in the Supporting Information online). The meaning of each word had to be selected from a choice of five options. The test revealed that the bilingual participants across both groups were highly variable in their performance, with scores ranging from 2 to 26 out of 30. The items selected for the vocabulary test were all quite difficult relative to the words selected for use in the lexical decision experiment and, therefore, even the participant who scored only 2 out of 30 should have been capable of performing the experiment. The alpha coefficient for the vocabulary test was .82, suggesting the items had high internal consistency. The vocabulary score was taken as a rough indication of English proficiency in that greater experience and use of English are likely to lead to increased word knowledge. Consistent with this assumption, the difference between the vocabulary scores of the two bilingual groups was highly significant, $t(99) = 5.84$, $p < .001$, $d = 1.19$. Moreover, vocabulary scores were correlated with the estimated weekly use of English, $r = .39, p < .001$, which is a medium-strength association according to Plonsky and Oswald's (2014) L2-specific guidelines.

Materials and Design

There were 22 items in each of the three conditions (see Appendix S2 in the Supporting Information online). The transparent word pairs possessed a semantic relationship with their stems (e.g., *hunter-HUNT*), while this was not the case in the opaque condition, despite the fact that the prime had the appearance of a stem plus suffix (e.g., *corner-CORN*). Pairs in the form condition were related orthographically but not semantically or morphologically (e.g., *freeze-FREE*). Primes in this condition comprised the target plus a nonmorphological ending (i.e., letters not used as a suffix in English, such as *–ze*). Log frequencies of both the primes and targets (taken from Davis, 2005) were matched across conditions, as were those for length and orthographic neighborhood size (i.e., the number of other words that differed from the word in question by one substituted letter). For each target, an unrelated prime was also selected, which was orthographically, morphologically, and semantically unrelated to its corresponding target. The unrelated primes were matched as closely as possible to each related prime on frequency and length.

To ensure that all participants were likely to be familiar with the targets, the materials were chosen on the basis of a set of previously collected familiarity ratings. A group of students from ZJU who did not participate in the test were asked to evaluate the familiarity of each item (including unrelated primes) based on a 7-point scale with 0 meaning "unknown" and 7 meaning "very familiar." Items with an average score lower than 5 were eliminated in order to optimize the familiarity of the items for bilinguals at the university level. Thus 132 word pairs were chosen. The mean values of all controlled item properties are shown in Table 2.

In addition to the 66 word targets, the experiment also included 66 nonword distractors, which all had real word primes, half of which were related to their primes by having only one letter different from the stem of the prime (e.g., *widely-WODE*, *organic-ORBAN*, *window-WINP*). Like the word–word pairs, two thirds of the word primes with nonword targets had a derivational ending (e.g., *–ly*, *–ic*), while a third had a nonmorphological ending (e.g., *ow*). The unrelated pairs with nonword targets were derived words or monomorphemic words that were morphologically or orthographically unrelated to the target (e.g., *timidly-WODE*, *wishful-ORBAN*, *perhaps-WINP*).

The targets were arranged in a Latin square design requiring two lists, each containing 22 pairs in each of the semantic transparent, opaque, and form-related conditions, with half of the items preceded by a related prime and the other half preceded by an unrelated prime. The nonword distractors also fell into a Latin square design. Participants were only presented with one of the two lists and therefore saw the target words only once. The same set of 10 practice items preceded the experimental trials in each of the two lists.

Procedure

DMDX software (Forster & Forster, 2003) was used to present the stimuli and record the data. Following the standard procedure adopted in the masked priming paradigm (see, e.g., Forster & Davis, 1984), each trial consisted of a 500-millisecond forward mask (########) which was replaced by the lowercase prime, followed 50 milliseconds later by the uppercase target. The target remained onscreen for 1,500 milliseconds or until a response was given, with an inter-trial interval of 1,000 milliseconds. Participants were tested in a quiet room. They were told in their native language that a series of letter-strings would appear on the screen and that they should press the right shift key if it was a real English word or the left shift key if it was not. They were instructed to respond as quickly and as accurately as possible. The participants were not told of the existence of the prime stimuli. When directly questioned at the end of the experiment, no participant said that s/he had seen any other letter string prior to the target item.

Table 2 Attributes of materials at different levels of prime by type

Word pair prime type			WL	ONS	LogF	Familiarity
			M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Transparent	Prime	Related	6.59 (1.01)	2.00 (1.72)	1.12 (0.33)	6.66 (0.24)
		Unrelated	6.59 (1.01)	2.68 (4.47)	1.13 (0.39)	6.56 (0.31)
	Target		4.41 (0.80)	9.91 (5.60)	1.59 (0.38)	6.64 (0.26)
Opaque	Prime	Related	6.68 (1.29)	2.91 (2.33)	1.40 (0.56)	6.45 (0.54)
		Unrelated	6.62 (1.18)	2.14 (2.05)	1.42 (0.65)	6.14 (1.79)
	Target		4.36 (1.18)	11.14 (6.37)	1.63 (0.56)	6.47 (0.40)
Form	Prime	Related	6.55 (1.30)	3.23 (3.02)	1.24 (0.65)	6.29 (0.71)
		Unrelated	6.59 (1.10)	1.95 (2.24)	1.22 (0.56)	6.65 (0.37)
	Target		4.18 (0.80)	11.32 (7.98)	1.68 (0.60)	6.49 (0.42)

Note. WL = word length; ONS = orthographic neighbourhood size; LogF = log frequency. Because target words are basic words with no suffix, they are necessarily shorter than prime words and also have more one-letter different neighbors.

After completion of the lexical decision task, bilingual participants were administered the vocabulary test, which they were asked to complete as quickly as possible. No dictionary or access to the Internet was allowed.

Results

As is now standard for lexical decision experiments, a linear mixed-effects analysis was conducted using R and the lme4 package (Bates, Maechler, & Bolker, 2012), with a focus on reaction times to correct responses and error rates. Before fitting the models, 4 out of the 101 bilingual participants were removed due to high error rates (i.e., higher than 45%, which was 2.5 standard deviations above mean group accuracy). For the rest, reaction times for incorrect responses were excluded from the analysis of response times, as were reaction times faster than 200 milliseconds and slower than 1,500 milliseconds, as is often the procedure when analyzing lexical decision responses (e.g., Heyer & Clahsen, 2015). The discarded data accounted for 4.72% of all the bilingual data and 2.29% of the monolingual data. Table 3 displays the mean reaction times for correct responses and percentage error rates from each condition for the remaining data.

The reaction times for correct responses were inversely transformed (i.e., –1,000/reaction time) to reduce a positive skew in the data. The transparency of the prime–target pair (transparent vs. opaque vs. form related), the group factor (monolingual vs. bilingual Group I vs. bilingual Group C), and prime type (related vs. unrelated) were treated as fixed effects. The length, frequency, and neighborhood size of the targets were centered around the mean (to decrease multicollinearity between any interaction term and its corresponding main effects) and, if significantly improving the models, were included as covariates. The analysis of reaction times also included the inverse-transformed response time on the previous trial as a fixed factor. A maximal random-effects model was used initially, and factors were removed successively based on likelihood ratio tests. The best fitted models had random intercepts by subjects and items as well as random slopes for conditions (by subjects).

The error rate analyses were conducted using mixed-effects logistic regression, with a binomial link function. Whether an error had been made in the previous trial was included as a fixed factor. Coefficients with two-tailed p values less than .05 (based on the z distribution) were considered statistically significant and the values of t and z are reported. In order to see whether an effect for each prime type existed for the two bilingual groups and the monolingual group, comparisons regarding the effect of relatedness for each prime type were conducted within each of the groups and in pairwise group comparisons

Table 3 Mean reaction time and error rate performance (standard deviations) for participant groups by prime type in Experiment 1

Word pair prime type		Reaction time (ms)			Error rate (%)		
		Group C	Group I	Monolinguals	Group C	Group I	Monolinguals
Transparent	Related	818 (243)	802 (234)	560 (136)	3.45 (18.26)	4.08 (19.81	4.17 (3.46)
	Unrelated	875 (236)	862 (230)	597 (115)	9.03 (28.68)	5.94 (23.66)	4.86 (21.53)
	Priming	57	59	37	5.58	1.86	0.69
Standardized coefficients[a]		0.18	0.16	0.18	1.81	0.84	0.73
Opaque	Related	867 (258)	836 (225)	594 (141)	10.84 (31.12)	7.32 (26.07)	5.50 (22.82)
	Unrelated	898 (244)	861 (211)	622 (153)	8.26 (27.55)	13.11 (33.79)	7.36 (36.14)
	Priming	31	26	28	–2.58	5.79	1.86
Standardized coefficients[a]		0.09	0.10	0.09	–0.42	1.15	0.63
Form	Related	832 (241)	822 (223)	607 (143)	7.77 (26.80)	6.58 (24.81)	7.95 (27.09)
	Unrelated	886 (228)	844 (205)	619 (147)	11.85 (32.34)	10.08 (30.27)	7.36 (26.14)
	Priming	54	22	12	4.08	3.5	–0.59
Standardized coefficients[a]		0.14	0.08	0.07	1.26	1.03	–0.14

Note. [a]Standardized coefficients are reported as an indication of effect size. They are calculated by multiplying the unstandardized standardized coefficient by the ratio of the standard deviations of the independent and the dependent variables.

(see Appendix S3 in the Supporting Information online for the full statistical output).

Reaction Time Analyses

When the reaction time data were analyzed across the three groups, there were main effects of relatedness, $t = 5.12$, $p < .001$; of group, $t = 11.63$, $p < .001$; and of target frequency, $t = 4.35$, $p < .001$, as well as a three-way interaction involving group, prime type, and relatedness, $t = 2.08$, $p = .04$.

For the monolingual group, the reaction time analysis showed a strong priming effect in both the transparent and the opaque conditions, $t = 5.24$, $p < .001$, $t = 2.46$, $p = .02$, respectively, but not in the form condition, $t = 1.67$, $p = .10$. There was a significant interaction between the magnitude of priming from the transparent and form-related conditions, $t = 3.02$, $p = .003$, and a marginal effect between transparent and opaque condition, $t = 1.99$, $p = .051$, but no interaction between opaque and form condition, $t = 0.80$, $p = .43$. The main effect of target frequency was also significant, $t = 3.73$, $p < .001$.

Similar to the monolingual group, bilingual Group I (international students) showed a strong facilitating effect in both the transparent, $t = 4.47$, $p < .001$, and opaque conditions, $t = 2.23$, $p = .03$, but no significant form priming, $t = 1.65$, $p = .10$. However, no significant interactions with priming were revealed across the three conditions except for a marginal effect when the transparent condition was compared with the form condition, $t = 1.79$, $p = .08$. For bilingual Group C (students in China), significant priming effects were seen in both the transparent and form conditions, $t = 3.44$, $p = .001$, $t = 2.32$, $p = .02$, respectively, but not in the opaque condition, $t = 1.45$, $p = .15$. However, there were no significant interactions between any of the three conditions (all *p*s $> .13$).

There were no interaction effects when Group I was compared with either the monolinguals or with Group C (all *ps* $> .12$). In the comparison of Group C and the monolinguals, the reaction time data showed a significantly stronger form- priming effect for the former than for the latter, $t = 2.29$, $p = .02$, while there was no interaction for the other two conditions: transparent items, $t = 1.20$, $p = .23$, and opaque items, $t = 0.35$, $p = .73$. There was also a significant interaction when the magnitude of transparent priming was compared with that of form priming between Group C and the monolinguals, $t = 2.07$, $p = .04$, because monolinguals showed greater priming in the transparent condition than in the form condition while Group C showed equivalent priming in these two conditions. No other significant interactions were observed across any other conditions (all *ps* $> .12$).

Error Rate Analyses

When the analysis of error rates was conducted across the three groups, a main effect of relatedness was found, $z = 2.01$, $p = .04$, as well as a significant target frequency effect, $z = 2.02$, $p = .04$. There was also a three-way interaction between relatedness, condition, and group, $z = 2.51$, $p = .01$.

The analysis of error rates for the monolingual group found no significant simple effects or interactions across any of the conditions (all ps $> .32$). For bilingual Group I, significant priming effects were found for the error rate measure in the transparent condition, $z = 2.15$, $p = .03$; the opaque condition, $z = 3.16$, $p = .002$; and the form condition, $z = 2.06$ $p = .04$, with no significant interactions across the conditions (ps $> .21$). For bilingual Group C, the error rate analysis showed a similar pattern to the reaction time analysis, with a significant transparent priming effect, $z = 2.51$, $p = .01$, and form-priming effect, $z = 2.36$, $p = .02$, but no sign of an opaque priming effect which, if anything, went in the reverse direction, $z = 0.78$, $p = .44$. There was a significant interaction with relatedness when the transparent condition was compared with the opaque condition, $z = 2.13$, $p = .03$, and also when the opaque condition was compared with the form condition, $z = 2.09$, $p = .04$.

There were no significant interactions on the error rate measure (ps $> .15$) when Group I was compared with the monolingual group. However, when Group C was compared with the monolingual group, a significant interaction was revealed in the opaque condition, $z = 2.02$, $p = .04$, which was also true when Group C was compared with Group I, $z = 3.58$, $p < .01$.

Discussion

The results of Experiment 1 revealed that both transparent and opaque words generated significant priming effects for monolingual participants.[2] In line with findings previously reported in the literature (e.g., Rastle et al., 2004), this can be interpreted as a morphological effect, given that there was no priming for form-related words (though such a conclusion is somewhat tempered by the fact that the interaction between opaque and form priming did not reach statistical significance). The performance of bilingual Group I was quite similar to that of the monolingual group, showing significant transparent and opaque priming but no form priming. Such a result is consistent with that of Diependaele et al. (2011). In contrast, bilingual Group C, who had less exposure to English than Group I and had a lower mean vocabulary score, showed significant transparent and form priming, but opaque priming was not statistically significant. This pattern was similar to the result of Zhang et al. (2016). The fact that there was no interaction between Group I and either Group C or the monolingual group suggests that the pattern of performance for Group I shared characteristics with both the less-proficient bilinguals and the native speakers.

The clear priming effects that bilinguals of lower proficiency (Group C) produced in the form-related condition, in both the reaction time and error rate measures, suggest that general orthographic factors exert influence at the early stages of processing. Form priming was stronger for this group of bilinguals than for monolinguals, and there was no difference compared to the priming observed in the transparent condition. Such a result is consistent with the findings of Heyer and Clahsen (2015), where bilinguals showed the same magnitude of form-based and morphological priming, while natives showed morphological priming but no form-based priming. Heyer and Clahsen argued that nonnative language processing is not necessarily less automatic than L1 processing but that early word recognition processes in a nonnative language are driven more by surface form properties than by morphological relatedness. The finding of the statistically larger priming effect in the form condition for Group C, compared to the monolingual group, suggested that Group C showed the strongest form priming, which indicated that less proficient L2 learners might rely more on orthographic processing.

Group C showed no opaque priming in either the reaction time or error rate measures, while Group I did. The performance of Group C was consistent, to some extent, with that found by Zhang et al. (2016), where bilinguals showed priming effects for the transparent condition, but not for the opaque condition when the stimulus onset asynchrony was 48 milliseconds. Two different accounts for the sequence of morpho-orthographic and morphosemantic activation at the early stage of visual word recognition have been formulated. The *form-then-meaning* account (e.g., Rastle & Davis, 2008; Rastle et al., 2004) asserts that the form of the word is processed before its meaning, whereas the *form-with-meaning* account (e.g., Feldman, O'Connor, & del Prado Martín, 2009) asserts that both orthographic and semantic features are activated at the same time. Zhang et al. claimed that their data supported the latter because transparent priming appeared earlier than opaque priming in their study. Thus, they argued for parallel activation of orthography and semantics in L2 decompositional processing, something that would only be possible in a cascaded system where semantic information is extracted from lexical candidates that have only been partially activated on the basis of form information.

Given that the low-proficiency bilinguals showed equivalent priming in the transparent and form conditions, it is possible that such individuals are only sensitive to orthographic information, with neither morpho-orthographic nor morphosemantic decomposition being applied. However, it is then very hard to explain why they showed no significant opaque priming when that condition was analyzed on its own. One suggestion might be that the competition that arises between *corner* and *corn* is inhibitory, counterbalancing the facilitatory effect of orthographic overlap that is seen in the form condition. However, this is very unlikely because it would mean that a group who is most sensitive to the fact that *corner* looks like a derived word but is not would be the group who is least proficient in English and shows no evidence of morphological decomposition when the transparent and form conditions are compared. Further research therefore needs to be carried out to explore the lack of opaque priming relative to form priming for low-proficiency bilinguals.

Masked priming effects adopted in Experiment 1 are supposed to reflect the early stages of visual word processing, showing the impact of the property of words on recognition before they are consciously processed. If the priming effects from Group C were a product of form overlap alone, it might mean either that these participants were ignorant of the morphological structure of derived English words or that they had a sensitivity to English morphology but did not activate those representations for such words via morpho-orthographic decomposition. These two possibilities are examined in Experiment 2, where a basic lexical decision task was employed to investigate how morphological structures are analyzed when novel words are presented.

Experiment 2

Morphological sensitivity is most clearly observed by looking at pseudowords that have a morphological structure. It could be argued that any morphological effects observed with pseudowords arise from processes that only come into play when whole-word access fails (e.g., Casalis et al., 2015; Taft, Hambly, & Kinoshita, 1986). If a person is sensitive to morphological structure, this should at least be observed for pseudowords (e.g., Burani & Laudanna, 2003; Longtin & Meunier, 2005; Marsden, Altmann, & St Claire, 2013; Taft et al., 1986). For example, Taft et al. (1986) found that nonexisting words composed of a mismatched prefix and stem (e.g., *dejoice*) were harder to reject in a lexical decision task than nonwords composed of a stem combined with a nonprefix (e.g., *tejoice*). Such results imply the existence of a decomposition mechanism that at least comes into play when pseudowords are processed.

In Experiment 2, bilinguals were tested to investigate whether they would treat pseudowords that were generated by combining a mismatched stem and suffix in the same way as monolinguals had done in a lexical decision task based on the following hypotheses. If bilinguals are simply insensitive to

Table 4 Bilingual participants' background variables and vocabulary knowledge in Experiment 2

	Group C	Group I
Variable	*M* (*SD*)	*M* (*SD*)
Age	19.1 (0.9)	23.1 (2.8)
Age of acquisition	9.6 (1.8)	9.7 (2.0)
Years of learning English	9.7 (1.5)	11.0 (1.9)
Weekly use of Chinese (%)	95.6 (4.4)	61.1 (10.8)
Weekly use of English (%)	10.3 (11.6)	38.9 (20.6)
Vocabulary test score	10.3 (4.8)	14.2 (4.6)

the morphological structure of derived words, they will be likely to process morphologically structured pseudowords just as they would process any pseudoword and, hence, find the former no harder to reject as a word than the latter. On the other hand, if even low-proficiency bilinguals have morphological sensitivity, they will act in a way similar to the way monolinguals respond and will find it harder to reject a morphologically structured pseudoword (even though the bilinguals tested in Experiment 1 did not reveal any clear evidence for automatic decomposition). Such an outcome would be consistent with the experiment of Casalis et al. (2015), which showed that French-English bilinguals found it hard to reject derived pseudowords regardless of their L2 proficiency.

Method

Participants

Another 49 English monolinguals were recruited from UNSW, receiving course credit for their participation. There were 40 Chinese-English bilinguals from UNSW and 42 from ZJU based on the same requirements as in Experiment 1, and each received A$10 or ¥20, respectively. None had participated in Experiment 1. All bilinguals received the same vocabulary test and questionnaire as in Experiment 1. Mean scores are shown in Table 4. A series of *t* tests suggested there were no significant differences between the bilingual groups of Experiment 1 and Experiment 2 in any of the measures (all ps $> .05$). Bilinguals were divided into two subgroups labeled Group I and Group C using the same criteria as in Experiment 1. The two bilingual subgroups again differed in their vocabulary scores, $t(80) = 4.78, p < .001, d = 1.06$, and weekly use of English, $t(80) = 15.14, p < .001, d = 3.40$.

Materials and Design

Thirty pseudowords[3] were constructed by combining a real word with an existing suffix (e.g., *animalful*, *natureness*, *discussly*). They will hereafter be referred to as suffixed nonwords. Another 30 control nonwords were constructed from the suffixed nonword items by changing one letter of the suffix (e.g., *animalfil*, *naturenass*, *discussny*). Both neighborhood size and bigram frequency (taken from Davis, 2005) were matched across the control nonword and suffixed nonword conditions. Because of the similarity of the control and suffixed nonword pairs, a Latin square design was again adopted whereby half of the suffixed nonword items were assigned to List 1 and their control nonword counterparts to List 2, while the other half of the suffixed nonword items were assigned to List 2 and their control nonword counterparts to List 1. In this way, no participants saw both members of a control nonword–suffixed nonword pair, yet both were presented equally often across the two lists. Fifteen derived words (e.g., *painful*) were also included in Experiment 2 as distractors, having the same suffixes as the nonword targets. A further 15 words were either prefixed (e.g., *design*) or monomorphemic (e.g., *secretary*) words.

Procedure

A basic lexical decision task was used in Experiment 2 (i.e., without any primes) using the same equipment as in Experiment 1. Items were presented in lowercase for 1,500 milliseconds or until a response was made. Participants received the same instructions as in Experiment 1.

Results

The method of analysis applied in Experiment 1 was also applied in Experiment 2. One monolingual and two bilingual participants with error rates higher than 2.5 standard deviations above mean group accuracy were removed, together with reaction times faster than 200 milliseconds or slower than 1,500 milliseconds, which accounted for 2.8% and 3.5% for monolingual and bilingual participants, respectively. Responses to one item (*blinder*) were eliminated from the analysis due to the realization that it was a real word that might be known to some participants. Mean reaction times across the three conditions are presented in Table 5 (with the full statistical output provided in Appendix S4 in the Supporting Information online). The best linear mixed-effects models were fitted with subject and item added as random factors and the length of target and bigram frequency as covariates. Ending frequency was also added as a covariate that was determined by the number of words generated by the last four letters for each suffixed and control nonword item according to the Web

Table 5 Mean reaction time and error rate performance (standard deviations) for participant groups by target type in Experiment 2

	Reaction time (ms)			Error rates (%)		
Lexical decisions	Group C	Group I	Monolinguals	Group C	Group I	Monolinguals
Derived words	777 (233)	754 (201)	607 (169)	11.34 (31.72)	3.25 (17.74)	3.90 (19.37)
Suffixed nonwords	1,040 (215)	1,078 (282)	775 (237)	40.23 (49.07)	42.65 (49.50)	27.99 (44.89)
Control nonwords	1,034 (212)	1,009 (253)	715 (206)	17.97 (38.42)	20.14 (40.14)	9.31 (29.09)
Difference	6	69	60	22.26	22.51	18.68
Standardized coefficients[a]	0.05	0.12	0.17	1.45	1.46	2.74

Note. [a]Standardized coefficients are reported as an indication of effect size. They are calculated by multiplying the unstandardized standardized coefficient by the ratio of the standard deviations of the independent and the dependent variables.

site http://www.onelook.com. For letter strings (like *tion*) that appear in more than 100 words, a value of 100 was adopted. Random slopes for conditions (by subjects) were included in the models.

For the monolingual participants, the reaction time analysis showed that suffixed nonword items were harder to reject than their control nonword counterparts, $t = 6.00$, $p < .001$. In relation to the bilingual participants, the interaction between the two subgroups and the suffixed versus control nonword comparison was significant, $t = 2.33$, $p = .02$, which reflected the fact that the main effect for suffixed versus control nonwords was significant for Group I, $t = 3.84$, $p < .01$, but not for Group C, $t = 1.50$, $p = .13$. The nonsignificant condition effect (suffixed vs. control nonwords) for Group C also interacted with the significant effect for the monolingual group, $t = 2.25$, $p = .02$, and Group I, $t = 2.12$, $p = .03$, while the magnitude of the effect was the same for Group I as for the monolingual group, with no significant interaction, $t = 1.63$, $p = .10$.

There were no interaction effects among three groups in the error rate analysis (ps $> .09$) because the error rate (incorrectly classifying the pseudowords as words) in the suffixed nonword condition was higher than in the control nonword condition for all three groups: monolinguals, $z = 7.91$, $p < .01$; Group I, $z = 6.72$, $p < .01$; and Group C, $z = 8.37$, $p < .01$. The length of the target was also significant in all models for both reaction times and error rates (ps $< .05$).

Discussion

The mean reaction times and error rates for both bilingual groups and the monolingual group were higher in the suffixed nonword condition than the control nonword condition, showing that it was harder for all groups to reject pseudowords that ended in an existing suffix than pseudowords that did not.

Both the reaction time and error rate analyses indicated that Group I participants were sensitive to morphological structure, and their performance showed a similar pattern to that of the monolinguals. Group C participants showed that they were just as likely to make more errors on the suffixed as on the control nonword items as were Group I participants and the monolinguals. However, while the accuracy measure provided evidence that Group C, unlike Group I, had awareness of morphological structure, the reaction time measure showed no difference between the suffixed and control nonword conditions. Perhaps on those occasions when Group C participants accessed the stem and affix of the pseudoword through decomposition, they were more inclined to accept it immediately as a word than were those in Group I, who may have been more prepared to consider whether the stem and affix really did combine to form a real word. That is, Group I participants were more advanced in their knowledge of the word formation rules of English and thus spent more time processing such linguistic information in the suffixed nonword items. Consistent with this idea is the fact that Group I had a longer mean reaction time than did Group C for the suffixed nonword items, despite being faster for the control nonword items and real words (see Table 5).

The results of this experiment appear to be compatible with the findings of Casalis et al. (2015), suggesting that readers with an L1 Chinese background do not differ with regard to morphological sensitivity in English from readers with an L1 French background, despite the difference that exists in the morphological richness of those two L1s. However, without a direct comparison using the same set of materials, conclusions about the relative morphological sensitivity of bilinguals from different L1 backgrounds must remain tentative.

It should be noted that the derivational suffixes used in the suffixed nonword condition were, by their very nature, more frequent as letter combinations in English than the equivalent endings in the control nonword condition. As such, the relative difficulty observed when responding to suffixed nonword items might have arisen from sensitivity to the frequency of letter combinations rather than morphology per se. An attempt was made to control for the frequency of the final letter combinations by entering the ending frequency value as a covariate in the analysis. Indeed, the very strong correlation between this value and the experimental condition suggested that ending frequency may influence decompositional processing and, potentially, its learning. That is, it is likely that the morphemic status of a suffix is learned as a result of the fact that it recurs in a number of different words that are associated with shared functional information (see, e.g., Longtin & Meunier, 2005). Thus, it is very hard to disentangle morphemic status from orthographic frequency, given the intrinsic nature of morphemes. Sensitivity to orthographic recurrence is a requirement for sensitivity to morphology, and the accuracy results of Experiment 2 suggested that even less proficient L2 learners were sensitive to the former at the very least (see Marsden, Williams, & Liu, 2013, for a similar claim when examining the acquisition of an artificial language).

General Discussion

The aim of this study was to investigate the degree to which Chinese-English bilinguals are sensitive to derived morphological structures in English and the extent to which any sensitivity is related to English proficiency as determined by experience and use.

In Experiment 1, significant priming was found for monolinguals in the transparent and opaque conditions but not in the form-related condition. Such a pattern of results is consistent with previous studies (e.g., Rastle et al., 2004) and indicates that morphological decomposition occurs on the basis of the orthographic appearance of the letter string, prior to any semantic processing. When the bilingual group was split into the two subgroups according to their learning context at the time of testing, an international group with higher vocabulary scores and more exposure to English showed priming patterns more like those of the native speaker group, while the group living in China showed significant priming in the transparent and form conditions but not in the opaque condition. This has parallels in other studies that have compared morphological and orthographic priming in L2 (e.g., Diependaele et al., 2011) and in which significant orthographic facilitation has been observed, even if to a lesser degree than morphological facilitation.

The results of form priming are consistent with the findings of Heyer and Clahsen (2015), which suggested that lower-proficiency bilinguals are influenced more by orthographic information than by morphological structures in early word recognition. Such a conclusion might seem to conflict with the apparent support for decomposition that Silva and Clahsen (2008) reported from their finding that Chinese-English bilinguals showed significant priming of stems by derived words relative to an unrelated baseline (e.g., *neatness-NEAT* vs. *dark-NEAT*), even if weaker than identity priming (e.g., *neat-NEAT*). However, it is possible to interpret the results of Silva and Clahsen in terms of mere orthographic similarity between the prime and target. That is, identity primes were orthographically identical to the target, unrelated primes had no orthographic similarity to the target, and derived primes had orthographic information that was not found in the target. Such an explanation is quite compatible with the results of Experiment 1 showing an equal magnitude of priming between the morphologically transparent structured and form-related conditions for the lower-proficiency bilinguals. However, Silva and Clahsen failed to find any priming effects for inflected primes, which also overlap orthographically with the target, so an explanation in terms of form overlap cannot be the whole story.

The finding that higher-proficiency bilinguals showed similar transparent morphological priming to that shown by lower-proficiency bilinguals in Experiment 1 might be seen as being consistent with the research of Diependaele et al. (2011) and Duñabeitia et al. (2013), who found that proficiency levels did not modulate the priming pattern for derived words. However, that the lower-proficiency group showed form priming to be just as strong as transparent priming suggested that the apparent transparent priming for this group could actually be seen as being form-based priming. If that were the case, though, there should have been just as strong a priming effect in the opaque condition as in the other two conditions, yet there was not. In fact, it is very hard to explain why lower-proficiency bilinguals showed no opaque priming but did show both transparent and form priming. Given the fairly small number of items per condition (i.e., 11 per participant within the Latin square design), it is necessary to establish whether such an outcome can be replicated in future research with a larger number of items. The study of Zhang et al. (2016) did indicate that opaque priming might not emerge under masked priming conditions for Chinese-English bilinguals. However, to be comparable with the present research, both an interaction with proficiency and a masked form-priming effect would need to be shown, neither of which were examined in that study.

The results of Experiment 2 testing pseudowords support the notion that bilingual participants do have a sensitivity toward derivational morphology, even if this is not brought into play when recognizing real words. Because there is no whole-word access unit for a pseudoword, the only units that can be activated are orthographically similar words or component morphemes. If morphological sensitivity is present, a morphologically decomposable pseudoword (e.g., *animalful*) should be more difficult to reject than a nondecomposable pseudoword (e.g., *animalfil*) because the former should activate the morpheme units *animal* and *–ful*, which lead to the corresponding representations in lexical memory. Because these components cannot be combined to form a word and thus the result must be classified as a nonword, processing should cost more time and generate more errors in lexical decision, compared to pseudowords that are not fully decomposable into morphemic constituents (e.g., Laudanna, Cermele, & Caramazza, 1997; Taft, 2004; Taft et al., 1986). Consistent with the findings of Casalis et al. (2015) with French-English bilinguals, the Chinese-English bilinguals of this study, as a group, showed that the decomposable pseudowords were harder to reject than nondecomposable ones in both the reaction time and error rate analyses, which meant that they were able to analyze a novel word according to its component morphemes and thus demonstrated sensitivity to morphological structure.

The analyses in both experiments demonstrated that the two bilingual groups were differentially affected by the morphological manipulations. Group I participants manifested similar priming patterns to monolinguals, while Group C participants did not show significant opaque priming in either reaction time or error rate measures in Experiment 1 nor a strong difference between the suffixed and control nonword items in the reaction time measure in Experiment 2. Such a difference between Group I and Group C can presumably be attributed to differences in the participants' experience with English. As the questionnaires suggested, Group I spent more time immersed in an English-language environment with more opportunity to use English, and the significant difference in vocabulary scores was consistent with such immersion leading to differences in their knowledge of English.

In classroom instruction, Chinese-English bilinguals are required to learn the specific function that suffixes play. Exercises in adding a suffix to a stem to generate a new word are commonly used as tasks in L2 English instruction, and the explicit teaching of suffix function enables learners to generate the meaning of novel morphemic combinations if they know the meaning of the stem words. It is only once learners have developed sensitivity to morphological structure that they will be able to decompose the derived words. Just as Marsden, Altmann et al. (2013) found that instructional activities that focused more on morpheme features than on sentence meaning elicited more morphological priming, so focusing L2 learners' attention on input features is likely to be very helpful. Moreover, Marsden, Williams et al. (2013) found that orthographic recognition was better if attention had been focused on the function of the morphemes during training prior to the crossmodal priming test. Exposure to such educational practices might explain why lower-proficiency bilinguals demonstrated morphological sensitivity, as seen in the accuracy measure of Experiment 2 when making judgments about pseudowords, but did not demonstrate clear evidence of morphological decomposition in Experiment 1.

Conclusion

In two experiments, it was demonstrated that Chinese-English bilinguals who had been immersed in an English-speaking environment processed morphological complex stimuli in a qualitatively similar manner to monolingual English speakers. Those who had not experienced such an environment, with less exposure and lower vocabulary scores, also revealed some awareness of morphological structure (as shown by their error rates to affixed pseudowords) but appeared to be more attuned to orthographic than morphological factors when processing real words, as revealed by their high level of form-based priming relative to morpheme-based priming.

Final revised version accepted 15 March 2017

Notes

1 The term “monolinguals” is used for convenience to refer to native English speakers even though they may not necessarily speak only one language. However, if they do speak another language, it is not Mandarin.

2 The fact that there was a strong trend toward an interaction between priming and transparency is consistent with the claim made by Feldman et al. (2009) that transparent priming is stronger than opaque priming. However, as pointed out earlier, this does not obviate the importance of the fact that opaque priming was nevertheless observed.

3 The term “pseudoword” is meant to imply that an item has never previously been encountered. Occasionally, however, an item might be so well structured that it is actually attested in the literature (e.g., *enjoyer*). Such cases are nevertheless very unlikely to have been encountered by the participants.

References

Andrews, S., & Hersch, J. (2010). Lexical precision in skilled readers: Individual differences in masked neighbor priming. *Journal of Experimental Psychology: General*, *139*, 299–318. https://doi.org/10.1037/a0018366

Basnight-Brown, D., Chen, L., Hua, S., Kostić, A., & Feldman, L. (2007). Monolingual and bilingual recognition of regular and irregular English verbs: Sensitivity to form similarity varies with first language experience. *Journal of Memory and Language*, *57*, 65–80. https://doi.org/10.1016/j.jml.2007.03.001

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-0. Retrieved July 20, 2015; http://CRAN.R-project.org/package = lme4

Burani, C., & Laudanna, A. (2003). Morpheme-based lexical reading: Evidence from pseudo-word naming. In E. Assink & D. Sandra (Eds.), *Reading complex words: Cross-language studies* (pp. 241–264). Dordrecht, Netherlands: Kluwer. https://doi.org/10.1007/978-1-4757-3720-2_11

Casalis, S., Commissaire, E., & Duncan, L. G. (2015). Sensitivity to morpheme units in English as L2 word recognition. *Writing Systems Research*, *7*, 186–201. https://doi.org/10.1080/17586801.2014.976165

Clahsen, H., Felser, C., Neubauer, K., Sato, M., & Silva, R. (2010). Morphological structure in native and nonnative language processing. *Language Learning*, *60*, 21–43. https://doi.org/10.1111/j.1467-9922.2009.00550.x

Clahsen, H., & Neubauer, K. (2010). Morphology, frequency, and the processing of derived words in native and non-native speakers. *Lingua*, *120*, 2627–2637. https://doi.org/10.1016/j.lingua.2010.06.007

Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, *37*, 65–70. https://doi.org/10.3758/BF03206399

Diependaele, K., Duñabeitia, J. A., Morris, J., & Keuleers, E. (2011). Fast morphological effects in first and second language word recognition. *Journal of Memory and Language*, *64*, 344–358. https://doi.org/10.1016/j.jml.2011.01.003

Duñabeitia, J. A., Dimitropoulou, M., Morris, J., & Diependaele, K. (2013). The role of form in morphological priming: Evidence from bilinguals. *Language and Cognitive Processes*, *28*, 967–987. https://doi.org/10.1080/01690965.2012.713972

Feldman, L. B., O'Connor, P. A., & Moscoso del Prado Martín, F. (2009). Early morphological processing is morphosemantic and not simply morpho-orthographic: A violation of form-then-meaning accounts of word recognition. *Psychonomic Bulletin & Review*, *16*, 684–691. https://doi.org/10.3758/PBR.16.4.684

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 680–698. https://doi.org/10.1037/0278-7393.10.4.680

Forster, K. I., & Forster, J. (2003). DMDX: A windows display program with millisecond accuracy. *Behavioral Research Methods, Instruments & Computers*, *35*, 116–124. https://doi.org/10.3758/BF03195503

Heyer, V., & Clahsen, H. (2015). Late bilinguals see a scan in scanner AND in scandal: Dissecting formal overlap from morphological priming in the processing of derived words. *Bilingualism: Language and Cognition*, *18*, 543–550. https://doi.org/10.1017/S1366728914000662

Kuo, L., & Anderson, R. C. (2006). Morphological awareness and learning to read: A cross-linguistic perspective. *Educational Psychologist*, *41*, 161–180. https://doi.org/10.1207/s15326985ep4103_3

Laudanna, A., Cermele, A., & Caramazza, A. (1997). Morpho-lexical representations in naming. *Language and Cognitive Processes*, *12*, 49–66. https://doi.org/10.1080/016909697386907

Liang, L., & Chen, B. (2014). Processing morphologically complex words in second-language learners: The effect of proficiency. *Acta Psychologica*, *150*, 69–79. https://doi.org/10.1016/j.actpsy.2014.04.009

Longtin, C., & Meunier, F. (2005). Morphological decomposition in early visual word processing. *Journal of Memory and Language*, *53*, 26–41. https://doi.org/10.1016/j.jml.2005.02.008

Longtin, C. M., Segui, J., & Hallé, P. A. (2003). Morphological priming without morphological relationship. *Language and Cognitive Processes*, *18*, 313–334. https://doi.org/10.1080/01690960244000036

Marsden, E., Altmann, G., & St Claire, M. (2013). Priming of verb inflections in L1 and L2 French: A comparison of "redundant" versus "non-redundant" training conditions. *International Review of Applied Linguistics in Language Teaching*, *51*, 271–298. https://doi.org/10.1515/iral-2013-0012

Marsden, E. J., Williams, J., & Liu, X. (2013). Learning novel morphology: The role of meaning and orientation of attention at initial exposure. *Studies in Second Language Acquisition*, *35*, 1–36. https://doi.org/10.1017/S0272263113000296

Packard, J. L. (2000). *The morphology of Chinese: A linguistic and cognitive approach*. Cambridge, UK: Cambridge University Press.

Parodi, T., Schwartz, B. D., & Clahsen, H. (2004). On the L2 acquisition of the morphosyntax of German nominals. *Linguistics*, *42*, 669–705. https://doi.org/10.1515/ling.2004.022

Plonsky, L., & Oswald, F. L. (2014). How big is 'big'? Interpreting effect sizes in L2 research. *Language Learning*, *64*, 878–912. https://doi.org/10.1111/lang.12079

Portin, M., Lehtonen, M., Harrer, G., Wande, E., Niemi, J., & Laine, M. (2008). L1 effects on the processing of inflected nouns in L2. *Acta Psychologica*, *128*, 452–465. https://doi.org/10.1016/j.actpsy.2007.07.003

Portin, M., Lehtonen, M., & Laine, M. (2007). Processing of inflected nouns in late bilinguals. *Applied Psycholinguistics*, *28*, 135–156. https://doi.org/10.1017/S014271640607007X

Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes*, *23*, 942–971. https://doi.org/10.1080/01690960802069730

Rastle, K., Davis, M. H., Marslen-Wilson, W. D., & Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes*, *15*, 507–537. https://doi.org/10.1080/01690960050119689

Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review*, *11*, 1090–1098. https://doi.org/10.3758/BF03196742

Silva, R., & Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition*, *11*, 245–260. https://doi.org/10.1017/S1366728908003404

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *57*, 745–765. https://doi.org/10.1080/02724980343000477

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, *14*, 638–647. https://doi.org/10.1016/S0022-5371(75)80051-X

Taft, M., Hambly, G., & Kinoshita, S. (1986). Visual and auditory recognition of prefixed words. *The Quarterly Journal of Experimental Psychology*, *38*, 351–366. https://doi.org/10.1080/14640748608401603

Ullman, M. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*, 231–270. https://doi.org/10.1016/j.cognition.2003.10.008

Ullman, M. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory and practice* (pp. 141–178). Washington, DC: Georgetown University Press.

Vainio, S., Pajunen, A., & Hyönä, J. (2014). L1 and L2 word recognition in Finnish: Examining L1 effects on L2 processing of morphological complexity and morphophonological transparency. *Studies in Second Language Acquisition*, *36*, 133–162. https://doi.org/10.1017/S0272263113000478

Zhang, Q., Liang, L., Yao, P., Hu, S., & Chen, B. (2016). Parallel morpho-orthographic and morpho-semantic activation in processing second language morphologically complex words: Evidence from Chinese-English bilinguals. *International Journal of Bilingualism*, *21*, 291–305. https://doi.org/10.1177/1367006915624249

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1. Vocabulary Test Items and Self-Rating Questions.
Appendix S2. Critical Items for Experiments 1 and 2.
Appendix S3. Full Statistical Output for Experiment 1.
Appendix S4. Full Statistical Output for Experiment 2.