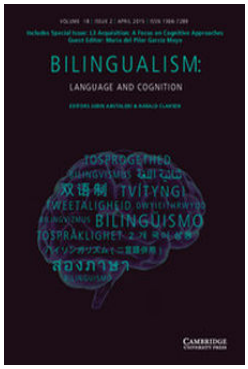
Bilingualism: Language and Cognition

http://journals.cambridge.org/BIL

Additional services for ***Bilingualism: Language and Cognition:***

Email alerts: Click here
Subscriptions: Click here
Commercial reprints: Click here
Terms of use : Click here



Cross language lexical priming extends to formulaic units: Evidence from eye-tracking suggests that this idea 'has legs'

GARETH CARROL and KATHY CONKLIN

Bilingualism: Language and Cognition / *FirstView* Article / April 2015, pp 1 - 19
DOI: 10.1017/S1366728915000103, Published online: 20 April 2015

Link to this article: http://journals.cambridge.org/abstract_S1366728915000103

How to cite this article:
GARETH CARROL and KATHY CONKLIN Cross language lexical priming extends to formulaic units: Evidence from eye-tracking suggests that this idea 'has legs'. Bilingualism: Language and Cognition, Available on CJO 2015 doi:10.1017/S1366728915000103

Request Permissions : Click here

CAMBRIDGE JOURNALS

Downloaded from http://journals.cambridge.org/BIL, IP address: 128.243.139.81 on 21 Apr 2015

CrossMark

Bilingualism: Language and Cognition: page 1 of 19 © Cambridge University Press 2015. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/3.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited. doi:10.1017/S1366728915000103

Cross language lexical priming extends to formulaic units: Evidence from eye-tracking suggests that this idea 'has legs'*

GARETH CARROL
University of Nottingham
KATHY CONKLIN
University of Nottingham

(Received: March 10, 2014; final revision received: February 24, 2015; accepted: February 25, 2015)

Idiom priming effects (faster processing compared to novel phrases) are generally robust in native speakers but not non-native speakers. This leads to the question of how idioms and other multiword units are represented and accessed in a first (L1) and second language (L2). We address this by investigating the processing of translated Chinese idioms to determine whether known L1 combinations show idiom priming effects in non-native speakers when encountered in the L2. In two eye-tracking experiments we compared reading times for idioms vs. control phrases (Experiment 1) and for figurative vs. literal uses of idioms (Experiment 2). Native speakers of Chinese showed recognition of the L1 form in the L2, but figurative meanings were read more slowly than literal meanings, suggesting that the non-compositional nature of idioms makes them problematic in a non-native language. We discuss the results as they relate to crosslinguistic priming at the multiword level.

Keywords: bilingualism, dual route processing, formulaic language, idioms, crosslinguistic influence

Introduction

Multi-word units (idioms, collocations, lexical bundles) have become an important focus in psycholinguistics. They are ubiquitous (Erman & Warren, 2000), show phrase-level effects of frequency (Bannard & Matthews, 2008; Tremblay, Derwing, Libben & Westbury, 2011) and have a privileged processing status for native speakers (Wray, 2012). However, they do not fit neatly into a 'words and rules' approach to language, so how such 'formulaic' units are processed and stored is a key question when it comes to understanding the structure of the mental lexicon.

Research into the bilingual lexicon has routinely looked at the relationship between single words in a first language (L1) and second language (L2) (Chen & Ng, 1989; de Groot & Nas, 1991; Wang, 2007), but there is a relative paucity of research into how translation equivalence might scale up to formulaic units. Some investigations of crosslinguistic influence have revealed an inherent reluctance to translate idioms (e.g., Kellerman, 1977, 1983, 1986), but other studies have shown effects of positive transfer, interference and avoidance in L2 idiom production (Irujo, 1986, 1993; Laufer, 2000) and comprehension (Liontas, 2001; Charteris-Black, 2002), generally finding facilitation for congruent items (those that exist in both languages). More recently, investigations into the online processing of such items have shown how congruence reduces the disruption caused during code switches in idiomatic and literal sentences (Titone, Columbus, Whitford, Mercier & Libben, 2015), and demonstrated the facilitatory effect of congruence in judging L2 collocations to be acceptable (Wolter & Gyllstad, 2011, 2013). We aim to add to this literature by exploring how translations of idioms are treated by intermediate proficiency Chinese–English bilinguals. Are 'familiar' sequences from the L1 treated as such even when they are encountered in an unfamiliar form? In other words, is the idiom priming effect that is evident when monolingual speakers read familiar phrases replicated when L1 idioms are encountered in the L2? The answer to this will have important implications for our understanding of how formulaic units are represented in the mental lexicon and will help to elucidate within-language relationships (how words are jointly represented) and between-language relationships (how different forms are represented across languages), both for single words and larger units. Translated idioms, therefore, provide a novel and potentially fruitful way to explore formulaic language in bilinguals. We begin by reviewing the existing literature on monolingual and bilingual idiom processing.

* This research was supported by a Sir Francis Hill Postgraduate Scholarship awarded to the first author. We are grateful to Gerrit Jan Kootstra and Debra Titone and to two anonymous reviewers for their extremely helpful comments on an earlier draft of this article.

Address for correspondence: Gareth Carrol School of English The University of Nottingham University Park Nottingham NG7 2RD
gareth.carrol@nottingham.ac.uk
Supplementary material can be found online at http://dx.doi.org/10.1017/S1366728915000103

CAMBRIDGE JOURNALS
http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81

In native speakers the processing advantage for familiar phrases is well documented. Using a range of methodologies, it has been demonstrated that highly familiar idioms are processed more quickly than less familiar idioms or control phrases (Cacciari & Tabossi, 1988; Conklin & Schmitt, 2008; Libben & Titone, 2008; McGlone, Glucksberg & Cacciari, 1994; Rommers, Dijkstra & Bastiaansen, 2013; Schweigert, 1986, 1991; Schweigert & Moates, 1988; Siyanova-Chanturia, Conklin & Schmitt, 2011; Swinney & Cutler, 1979; Tabossi, Fanari & Wolf, 2009). This evidence supports hybrid models, whereby idioms exist in the mental lexicon both as individual words and whole units, variously described as CONFIGURATIONS (Cacciari & Tabossi, 1988), SUPERLEMMAS (Sprenger, Levelt & Kempen, 2006) or FORMULEMES (Van Lancker Sidtis, 2012). The view that frequently encountered combinations are lexicalised to instantiate their own unitary representations in the mental lexicon is consistent with usage based accounts of linguistic organisation (e.g., Bybee, 2006, 2008), and the processing of these lexicalised units and their component parts can be accounted for in different ways. Libben and Titone (2008; also Titone & Connine, 1999) describe a constraint-based view of idiom processing which utilises all possible information to help process any given combination of words appropriately; this helps to address the 'paradox' of idioms seeming to be simultaneously unitary and compositional (Smolka, Rabanus & Rösler, 2007, p. 228). Dual route explanations of the formulaic processing advantage (Van Lancker Sidtis, 2012; Wray, 2002; Wray & Perkins, 2000) propose that all linguistic material is analysed sequentially as it is encountered, but an additional (and quicker) direct route is also available for those sequences that have been encountered previously and registered as known combinations. Once an idiom or other formulaic sequence is triggered/recognised, it can therefore be accessed directly.

While this effect is robust in native speakers, second language learners rarely show the same level of formulaic advantage (Cieślicka, 2006, 2013; Conklin & Schmitt, 2008; Siyanova-Chanturia et al., 2011; although see Isobe, 2011 and Jiang & Nekrasova, 2007 for alternative views). Second language learners may exhibit a fundamentally more compositional approach whereby sequential analysis is the default, meaning that literal meanings of words are likely to be more salient than figurative phrase-level meanings (Cieślicka, Heredia & Olivares, 2014). The question is whether this is actually a difference in approach or simply in available resources: non-native speakers may not have encountered idioms in the L2 with enough regularity to allow for formation and direct retrieval of unitary entries. This is not to say that idioms cannot be understood in the L2, but the same direct processing route may not be available by default (or may be too slow to show any effect). The present investigation aims to explore this question by looking at combinations that are theoretically 'known' to non-native speakers, but which are encountered in an unfamiliar (translated) form. Given that congruence seems to facilitate L2 processing of formulaic language (Titone et al., 2015; Wolter & Gyllstad, 2011, 2013), it remains to be seen whether this is a direct effect of L1 knowledge. That is, are congruent forms facilitated because they have been encountered in both languages and are confirmed in the minds of bilinguals as transferrable, or is it the case that any lexical combinations that exist in the L1 will automatically show priming effects if the equivalent forms are encountered in an L2? For example, when a French–English bilingual speaker first encounters *bite the dust* (a word-for-word equivalent of the French *mordre la poussière*), will this automatically be treated as an idiom because the forms are congruent, or would it only be accepted once the English version has been registered as the same as in the L1? In the present study we aim to investigate this for idioms that exist in the L1 but not the L2 (e.g., *call a cat a cat* – a non-idiom in English but a translation of the French *appeler un chat un chat*). Such items are therefore imbalanced in their relative L1–L2 frequency, hence any evidence of facilitation would be indicative of direct L1 influence.

There is some evidence that idioms should be processed quickly in their translated forms. Carrol and Conklin (2014) used a primed lexical decision task to show that intermediate proficiency Chinese–English bilinguals responded more quickly to idiom targets than control targets for items translated from the L1. When shown a prime of *draw a snake and add* . . . (a translation of the Chinese 畫蛇添足 – *draw-snake-add-feet* = *draw a snake and add feet*, meaning "to ruin something by adding unnecessary detail"), Chinese native speakers responded more quickly to the idiom target *feet* than they did the control target *hair*, whereas English native speakers showed no difference. Interestingly, in a similar study with Japanese collocations, Wolter and Yamashita (2014) found no advantage for acceptable L1 items presented in L2, so the extent of the effect remains unclear. Carrol and Conklin (2014) proposed two possible mechanisms underlying their pattern of results. The first is a lexical/translation route whereby English words automatically activate Chinese equivalents. A number of studies (e.g., Thierry & Wu, 2007; Wu, Cristino, Leek & Thierry, 2013; Wu & Thierry, 2010; Zhang, van Heuven & Conklin, 2011) have demonstrated that bilingual language processing may be non-selective in this way. Thus, it is plausible that when bilinguals read the prime phrases in English, the Chinese translations were automatically activated as each word was encountered. A known character sequence in the L1 was therefore triggered, making the final character available and in

CAMBRIDGE JOURNALS
http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81

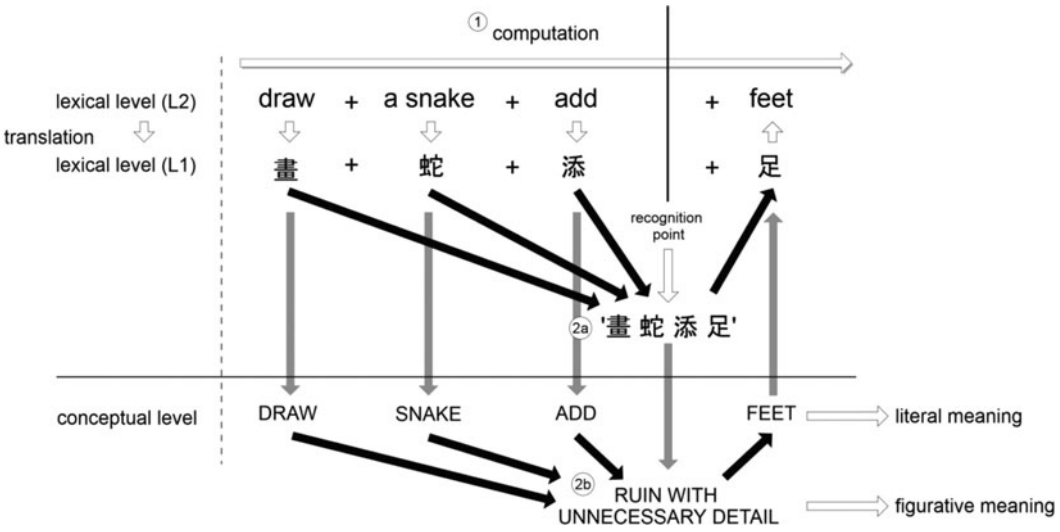


Figure 1. Dual route model of novel/familiar language processing, adapted to include translated idioms. A 'default' computation/analysis route is available (1), alongside two direct idiom retrieval mechanisms: a lexical-translation route (2a) and a conceptual priming route (2b). Black arrows represent associative links between components, white arrows represent processes and grey arrows represent links between lexical items and their underlying concepts. Reproduced from Carrol and Conklin (2014).

turn priming its translation equivalent in English. The second possibility is a conceptual route, whereby English (L2) words directly triggered their underlying concepts. The association of concepts (e.g., DRAW, SNAKE, ADD) triggered the underlying idiom concept (RUIN WITH UNNECCESSARY DETAIL), which activated the associated lexical components, either directly in the L2 if strong L2-conceptual links had been built up, or else in the L1, again priming the translation equivalent in English. This conceptual priming mechanism fits the suggestion by Wray (2012) that the advantage for idioms may be a result of their distinct underlying concepts.

Both mechanisms can be incorporated into the dual-route theory of familiar/novel language processing outlined in Figure 1.

The current research presents two experiments designed to explore idiom priming in bilingual speakers, using eye-tracking as a way to tap into the automatic processes at play during reading. The aim of Experiment 1 was to investigate whether the local lexical context provided by an idiom was enough to facilitate lexical access to the final word. We compared reading for idioms (*draw a snake and add feet*) and control items (*draw a snake and add hair*). Both variants were embedded in a short context that supported the idiomatic meaning, but neither would make sense in English without knowledge of the Chinese idiom. Shorter reading times for the final word in the idiom condition compared to the control would therefore be taken as evidence that bilingual speakers were utilising L1 knowledge to activate a known lexical combination and facilitate the expected completion.

The aim of Experiment 2 was to further explore the dimension of meaning in idiom processing. We specifically examined idioms that could also be used in a literal sense – what Van Lancker, Canter and Terbeek (1981) called 'ditropic' idioms. Hybrid models suggest that literal meaning activation is obligatory (Cacciari & Tabossi, 1988; Cieślicka & Heredia, 2011; Holsinger & Kaiser, 2010; Sprenger et al., 2006; but see Schweigert, 1991, on how relative familiarity and literal plausibility might moderate this). Siyanova-Chanturia et al. (2011) found that English native speakers showed comparable reading times for figurative and literal uses of highly familiar idioms: they read *at the end of the day* equally quickly in its idiomatic and literal senses, and both faster than a control phrase like *at the end of the war*. Non-native speakers read the literal uses significantly more quickly than the idiomatic uses, suggesting that the non-compositional nature of the figurative uses was problematic, or that the figurative meaning was simply not known. If L1 knowledge is being automatically activated when non-native speakers encounter translated forms, we would expect them to have little difficulty interpreting idioms in figurative contexts, hence we would expect the patterns of performance for Chinese native speakers on translated idioms to mirror that of English native speakers on English idioms, with no difference between figurative and literal uses for 'known' sequences.

In both experiments we compare Chinese native speakers and monolingual English native speakers reading translated Chinese idioms/controls and English idioms/controls.

CAMBRIDGE JOURNALS
http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81

Table 1. *Summary of Chinese native speakers' language background for both experiments (all measures relate to proficiency in English).*

	Age	Years studying English	Reading	Listening	Speaking	Writing	Usage	Vocab
Experiment 1								
Mean	26.2	14.1	3.4	3.2	2.8	2.8	35.3	10.7
Range	21–38	4–25	2–5	2–4	2–4	2–4	25–45	4–16
Experiment 2								
Mean	23.4	12.7	3.5	3.0	3.1	3.7	35.5	10.9
Range	21–30	7–16	2–5	1–4	1–4	2–5	27–43	7–16

N.B. Reading, Listening, Speaking and Writing are self-ratings of these skills out of 5 (1 = Poor, 2 = Basic, 3 = Good, 4 = Very good, 5 = Excellent); Usage is an aggregated estimate of how frequently participants use English in their everyday lives in a variety of contexts (total score out of 50); Vocab is a modified Vocabulary Size Test with a total score out of 20.

Experiment 1

In Experiment 1 we investigated whether 'known' sequences are facilitated in L2: do native speakers of Chinese show facilitation for the final word of a translated idiom compared to a control word? Chinese is ideal for this kind of investigation because it has a large set of invariable idioms (*chengyu*) that are numerous in modern Chinese. The vast majority are a fixed sequence of four characters[1] and *chengyu* have been shown to have the same formulaic properties as English idioms (Liu, Li, Shu, Zhang & Chen, 2010; Zhang, Yang, Gu & Ji, 2013; Zhou, Zhou & Chen, 2004).

Methodology

Participants

Participants in Experiments 1 and 2 were taken from the same population, but were different in each study. All participants received course credit or £5 for participation. Chinese native speakers were students at the University of Nottingham (34 postgraduates, seven undergraduates; mean age = 24.8), hence had met minimum entry requirements to study at an English university (minimum IELTS score of 6.5), and had been in the UK for an average of 1.4 years. All had Mandarin Chinese as their L1.[2] Information regarding their English language background is shown in Table 1. English native speakers were undergraduate students at the University of Nottingham (mean age = 19.3), none of whom had any experience of learning Mandarin. Twenty English native speakers and 20 Chinese native speakers took part in Experiment 1. All norming described below used participants who did not take part in the main experiments and used a seven-point rating scale.

Materials

Chinese idioms were selected from the Dictionary of 1000 Chinese Idioms (Lin & Leonard, 2012). Only common idioms where a literal translation provided a plausible English sequence with identical word order were considered, e.g., 畫蛇添足 – *draw-snake-add-feet* = *draw a snake and add feet*. For all items the final character had a single word translation equivalent in English. These idioms were judged to be highly familiar in the original Chinese form (mean = 6.5/7) by 27 native speakers of Mandarin. Translations were taken from the gloss provided by the Dictionary of 1000 Chinese Idioms then checked character by character using two different translation engines (Google Translate and Online Chinese Tools) to ensure accurate transliterations into English. Control items were formed by replacing the final word of each idiom with an alternative, matched for part of speech, length and frequency (e.g., *draw a snake and add feet* vs. *draw a snake and add hair*). All Chinese idioms and control items showed a phrase frequency of 0 in the British National Corpus (BNC). Note that the intention was not necessarily to create a literally plausible control sentence in each case, but simply to replace the final word in such a way that we could compare speed of access based on the preceding sequence. Hence in the example of *draw a snake and add feet/hair*, neither is inherently more literally plausible in English unless the idiom is

[1] Liu et al. (2010) estimated that around 97% of all *chengyu* in common use in modern Mandarin conform to the fixed four-character structure.

[2] Whilst the participants were all native speakers of Mandarin, it is possible that their time spent living in England could have led to slight language attrition that may have affected performance in these studies. The majority of students had spent a little over one year in England and when asked about their daily usage of Mandarin suggested that this was frequent since their social lives were largely constructed around other native speakers of Mandarin, although no data was collected to confirm this. We therefore assume that any effects of attrition would be minimal, but for future studies looking at L1 influence it might be useful to consider this in a more rigorous fashion.

CAMBRIDGE JOURNALS
http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81

known, but if Chinese native speakers are activating the underlying L1 idiom then this should lead to facilitation for the expected word.

English idioms were selected from the Oxford Learner's Dictionary of English Idioms (Warren, 1994). Twenty-six idioms were judged to be highly familiar (mean = 6.6/7) by 19 English native speakers. Control items were formed by replacing the final word with an alternative matched for part of speech, length and frequency (e.g., *spill the beans* vs. *spill the chips*). As with the Chinese items, the intention was not to create literally plausible control items but rather to specifically test whether the 'correct' word was facilitated once an idiom had been encountered. All control items showed a phrase frequency of 0 in the BNC. The English and Chinese items used in both experiments are available in Appendix S1 in the Supplementary Materials Online (Supplementary Materials).

All stimulus items were embedded in short sentence contexts supporting the figurative meaning, for example: "My wife is terrible at keeping secrets. She loves any opportunity she gets to meet up with her friends and *spill the beans/chips* about anything they can think to gossip about." All sentence contexts were of comparable length. Contexts for idioms and their corresponding controls were identical and all passages were presented over three lines with the idiom or control phrase appearing toward the middle of the second line. Forty filler items of comparable length were constructed, none of which contained idioms.

Compositionality ratings (how easily a literal paraphrase can be mapped onto an idiom) were gathered for all items, as this is often identified as an important factor in idiom processing (Caillies & Butcher, 2007; Gibbs, 1991; Gibbs, Nayak & Cutting, 1989). Sixteen English native speakers were presented with all English and Chinese idioms and asked how easily the meaning of the idiom could be matched to a literal equivalent (e.g., *to spill the beans* means "to reveal a secret"): English idioms: mean = 4.1/7; Chinese idioms: mean = 3.8/7. The Chinese idioms were also presented in the original Chinese characters to 12 Chinese native speakers who gave their own set of ratings (mean = 5.6/7).

Two counterbalanced stimulus lists were constructed so that each participant saw 13 English idioms, 13 English controls, 13 Chinese idioms, 13 Chinese controls and 40 filler items. Lists were matched for all lexical variables, for English idiom frequency and for the familiarity and compositionality of the idioms.

Procedure

The experiment was conducted using an Eyelink I (version 2.11) eye-tracker. Participants were seated in front of a monitor and fitted with a head-mounted camera to track pupil movements. Camera accuracy was verified using a nine-point calibration grid and recalibrations were performed throughout the experiment as required. Participants were asked to read the passages on screen for comprehension then press a button to advance once they had finished. Half of the items were followed by a yes/no comprehension question to encourage participants to pay attention and the rest were followed by a 'Ready?' prompt. After each trial a fixation dot appeared on the screen to allow for trial-by-trial drift correction. Each participant saw eight practice items, then the experiment began.

Afterwards, participants were asked to provide subjective familiarity ratings for all stimulus items. For English native speakers all items were presented in English (English items, mean = 6.4/7; Chinese items, mean = 2.1/7). For Chinese native speakers the English items were presented in the same way (mean = 3.5/7) but Chinese idioms were presented in the original Chinese characters (mean = 6.5/7).[3] Chinese native speakers were also asked to complete a short vocabulary test (modified from Nation & Beglar, 2007). This test was adapted to include a representative sample from the 10,000 most frequent word families in English, and was augmented with any low frequency vocabulary items that appeared in the stimulus items: for example, in the Chinese idiom *bare fangs and show claws, fangs* might be an unfamiliar English word, so we included such items in the test. Any constituent words from the English or Chinese idioms that were outside the 3000 most frequent word families in English were added to the test, and incorrectly identified words were removed from the analysis on a per-participant basis. Finally, Chinese native speakers were asked to complete a language background questionnaire (see Table 1 for details).

Analysis and Results

One Chinese native speaker was removed from the analysis because of eye-tracker calibration problems. All data were cleaned according to the four stage procedure within Eyelink Data Viewer software, meaning that fixations shorter than 100 ms and fixations longer than 800 ms were removed. Data were visually inspected and any trials where track loss occurred were removed, along with any trials containing words that were incorrectly identified on the vocabulary test (for non-native speakers only). Overall this accounted for 10.4% of raw data being removed from the analysis for Chinese native speakers.[4] No native speakers were removed from the analysis and 4.8% of the raw data was removed because of track

[3] It is worth noting that such high levels of familiarity with these Chinese character sequences perhaps argue against any significant level of attrition for the Chinese native speakers.

[4] Despite this relatively high figure, the composition of the stimulus lists in terms of key balancing factors like word length, frequency and idiom familiarity was not differentially affected for idioms vs. control items.

CAMBRIDGE JOURNALS
http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81

Table 2. *Summary of reading patterns of final words of phrases for all measures for Chinese native speakers and English native speakers.*

	Chinese phrases		English phrases	
	Idiom	Control	Idiom	Control
Chinese native speakers				
Likelihood of skipping	0.03	0.00	0.04	0.03
First fixation duration	272	301	269	262
First pass reading time	344	380	307	315
Total reading time	484	538	440	453
Total fixation count	1.8	1.9	1.7	1.7
English native speakers				
Likelihood of skipping	0.07	0.09	0.31	0.09
First fixation duration	199	201	134	183
First pass reading time	226	229	140	188
Total reading time	278	282	148	242
Total fixation count	1.3	1.3	0.8	1.2

Data are mean values for likelihood of skipping expressed as a probability, raw values in ms for duration measures and raw values for fixation counts. Mean duration measures include a value of zero for skipped items.

loss. Participants generally had no difficulty answering the comprehension questions (English native speakers, mean = 93%; Chinese native speakers, mean = 89%), suggesting that the task of reading and understanding the passages was well within the capability of all participants.

We concentrated the analysis on the final word of each phrase with the rationale that if idioms are known and stored as whole units then reading the first few words should activate the underlying phrase. This in turn should facilitate the final word relative to any other completion, and this would be reflected in shorter reading times. For items that are unknown we would expect to see no difference in reading times for an idiom vs. a control since no expectation regarding the final word would be generated. Although there was some variability in how literally plausible the phrases were, if an item was unknown to any participant then there should be no expectation generated for either the correct or incorrect ending.

We utilised a range of early and late eye-tracking measures to examine the predictability of the final word. Broadly, early measures reflect automatic lexical access processes while late measures reflect post-lexical processes/integration of overall meaning into wider context (c.f. Altarriba, Kroll, Sholl & Rayner, 1996; Inhoff, 1984; Paterson, Liversedge & Underwood, 1999; Staub & Rayner, 2007). Our early measures are probability of skipping (how likely is it that a word is not fixated during first pass reading), first fixation duration (duration of the first fixation on the final word of the phrase) and first pass reading time (sum of all fixations before gaze exited either to the left or right). The late measures are total reading time (sum of all fixations on the target word throughout any given trial, including re-reading time) and total number of fixations (total number of times a target word was fixated during any given trial). Table 2 shows a summary of the word-level reading patterns.

We analysed the data in an omnibus linear mixed effects model using the lme4 package (version 1.0–7, Bates, Maechler, Bolker, Walker, Christensen, Singmann & Dai, 2014) in R (version 3.1.2, R Core Team, 2014). Linear mixed effects models are able to incorporate random variation by subject and by item alongside fixed effects, thereby avoiding the "language as fixed effect fallacy" (Clark, 1973). We included the three treatment-coded main effects of group (Chinese native speakers vs. English native speakers), language (Chinese phrases vs. English phrases) and phrase type (idiom vs. control). Random intercepts for subject and item and by-subject random slopes for the effects of language and type were included (Barr, Levy, Scheepers & Tily, 2013). We included the covariates of idiom length in words, final word length in letters and log-transformed final word frequency in a stepwise fashion and compared the resulting models using likelihood ratio tests to see whether inclusion improved the fit; only covariates that significantly improved the model were retained. Separate models were fitted for each eye-tracking measurement. For the binary measure likelihood of skipping a logistic linear model was used (Jaeger, 2008). For subsequent analysis of durational

CAMBRIDGE JOURNALS
http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81

measures any skipped items were removed from the dataset and all duration measures were log-transformed to reduce skewing. Fixation counts were analysed using a generalised linear model with poisson regression. The structure and output for all models is shown in Table 3.

In an initial model for skipping rates there was a significant three-way interaction of group, language and type ($z = -2.63$, $p < .01$). English native speakers showed a strong tendency to skip the final word of English idioms (31%) compared to control items (9%) but no effect for Chinese items. Chinese native speakers showed a small but non-significant tendency to skip the final words of translated idioms vs. controls and no difference for English items. The analysis of duration measures also supports a general pattern whereby L1 idioms are read more quickly than control words: native speakers of Chinese read the final word more quickly for translated idioms vs. controls but show no difference for English phrases, while English native speakers show an advantage for English idioms but not translations of Chinese phrases. This is seen in the three way interaction of group, language and type: for first fixation duration this is marginal ($t = 1.84$, $p = .07$) and is significant for total reading time ($t = 2.31$, $p < .05$) and fixation count ($t = 3.33$, $p < .001$). For first pass reading time this interaction is not significant, but it must be remembered that this analysis has excluded all data for which the final word was skipped, which affected significantly more idioms than control phrases.[5]

Interactions were analysed further using the Phia package (version 0.1–5, De Rosario-Martinez, 2013) in R with separate models for the two speaker groups (available in Appendix S2, tables S2–3, Supplementary Materials). Pairwise comparisons confirmed that Chinese native speakers showed an advantage for Chinese idioms vs controls for first fixation duration (χ^2 (1, 841) = 5.39, $p < .05$), total reading time χ^2 (1, 841) = 4.81, $p = .05$) and marginally for first pass reading time (χ^2 (1, 841) = 4.12, $p = .08$), but not for likelihood of skipping or fixation count. For English phrases no differences were significant. English native speakers showed significantly higher likelihood of skipping for English idioms vs controls (χ^2 (1, 990) = 29.30, $p < .001$), significantly shorter total reading times (χ^2 (1, 990) = 5.78, $p < .05$) and significantly fewer fixations overall (χ^2 (1, 990) = 19.70, $p < .001$), but early duration measures were non-significant (again, most likely because of the high number of idioms that were removed from durational analysis because the final word was skipped). Chinese phrases showed no difference on any measure.

Phrase-level patterns

We also examined phrase-level data to see whether the overall context could have contributed to the pattern described above. We considered first pass reading time, total reading time (including re-reading) and regression path duration for the phrase (once the phrase had been fixated, how much time was spent re-reading the context that preceded it). We also considered regression path duration specifically for the final word. These measures are summarised in Table 4.

The omnibus analysis (Table 5) shows significant interactions of group and language for all measures and a significant three way interaction for group, language and type (for all measures except phrase-level regression durations). English native speakers had a tendency to read English idioms faster and to regress less. For control items, encountering an unexpected final word caused a regression to the immediate preceding context, but there was no difference in the amount of time spent re-reading the context prior to the phrase for idioms vs. controls. There was no difference between Chinese idioms and controls on any measure.

Chinese native speakers showed no difference on any of the phrase-level measures for idioms compared to controls for either set of phrases (pairwise analysis by type, all ps $> .05$). Encountering the 'incorrect' completion of an idiom from either language did not lead to more time re-reading the phrase. Similarly, whole phrase reading times and overall regressions to the preceding context were comparable for both sets of idioms and controls. One way to interpret this is that the recognition of form (evidenced in the analysis of the final words) and integration of meaning may be exerting opposing forces. That is, Chinese native speakers may be reading the idiom and correctly predicting the final word, but they still need to spend time reading and re-reading the whole phrase and the prior context to attempt to resolve the meaning in both idiom and control conditions. This hints at a dissociation between recognition/prediction of the correct form and access to the overall phrase-level meaning, which we will explore in more detail in Experiment 2.

Familiarity, Compositionality, Plausibility

We next analysed the data to assess the effect of subjective familiarity, relative compositionality and plausibility on each set of idioms. One possibility is that the difference in plausibility between idioms and controls might be exerting an effect: hence the advantage observed for idioms may in fact be a reflection of the disruption caused by implausible completions in the control items. To investigate this we collected plausibility ratings from 19 English native speakers to compare idioms and controls

[5] With this in mind, we also conducted a separate durational analysis where we retained all items but assigned all skipped words a single fixation duration of 100 ms (the lower cut off in our dataset). This analysis revealed highly significant three way interactions for all measures (duration measures, all ts>3, all ps $< .01$; fixation count, $t>2$, $p < .05$). See Appendix S2, table S1 (Supplementary Materials) for the full output of this model.

CAMBRIDGE JOURNALS
http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81

Table 3. *Omnibus linear mixed effects model output for final word, all eye-tracking measurements.*

	Likelihood of skipping			First fixation duration			First pass reading time			Total reading time			Total fixation Count		
Fixed effects:	β	SE	z	β	SE	t	β	SE	t	β	SE	t	β	SE	z
Intercept	−0.85	0.75	−1.14	5.55	0.03	164.09***	5.85	0.01	46.88***	5.83	0.11	53.66***	0.24	0.13	1.88
Group: English	0.85	0.50	1.71	−0.24	0.05	−5.43***	−0.34	0.06	−6.06***	−0.44	0.08	−5.81***	−0.28	0.09	−2.88**
Language: English	−0.07	0.58	−0.12	0.02	0.04	0.57	−0.02	0.04	−0.61	−0.04	0.06	−0.70	−0.02	0.08	−0.25
Type: Control	−2.03	1.12	−1.81	0.08	0.03	2.46*	0.09	0.04	2.25*	0.12	0.05	2.31*	0.08	0.07	1.09
Group*Language	1.64	0.65	2.53*	−0.09	0.05	−1.83	−0.08	0.05	−1.45	−0.21	0.07	−3.17**	−0.49	0.12	−4.17***
Group*Type	2.24	1.17	1.91+	−0.05	0.05	−1.10	−0.07	0.06	−1.16	−0.10	0.07	−1.39	−0.10	0.11	−0.93
Language*Type	1.18	1.23	0.96	−0.13	0.05	−2.90**	−0.09	0.05	−1.80	−0.10	0.06	−1.48	−0.08	0.10	−0.78
Group*Language*Type	−3.42	1.30	−2.63**	0.12	0.06	1.84	0.08	0.07	1.09	0.21	0.09	2.31*	0.53	0.16	3.33***
Control predictors:															
Word length (letters)	−0.58	0.13	−4.40***	n/a	n/a	n/a	0.02	0.01	1.77	0.03	0.02	1.88+	0.06	0.02	2.90**
LogFrequency	n/a	n/a	n/a	n/a	n/a	n/a	−0.03	0.00	−2.86**	n/a	n/a	n/a	n/a	n/a	n/a
Random effects	Variance			Variance			Variance			Variance			Variance		
Subject	0.000			0.011			0.019			0.037			0.033		
Item	0.176			0.001			0.002			0.011			0.006		
Subject | Language	0.254			0.003			0.001			0.000			0.001		
Subject | Type	0.461			0.002			0.010			0.012			0.001		
Residual	n/a			0.106			0.141			0.217			n/a		

Significance values are estimated by the R package lmerTest (version 2.0–11; Kuznetsova, Brockhoff & Christensen, 2014): *** $p < .001$, ** $p < .01$, * $p < .05$, + $p < .10$
For likelihood of skipping a logistic linear mixed effects model was used and for fixation count a generalised linear model with poisson regression was used.

CAMBRIDGE JOURNALS
http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81

Table 4. *Phrase-level reading patterns (all values in ms) for Chinese and English native speakers, all items.*

	Chinese phrases		English phrases	
	Idiom	Control	Idiom	Control
Chinese native speakers				
First pass reading time	1397	1411	904	915
Total reading time	1959	2030	1348	1382
Regression duration (word)	748	850	674	703
Regression duration (phrase)	1213	1180	680	683
English native speakers				
First pass reading time	814	736	423	482
Total reading time	1244	1238	528	681
Regression duration (word)	511	495	199	354
Regression duration (phrase)	745	746	334	352

for both English and Chinese phrases. English phrases were considered more plausible than the controls (idioms: mean = 6.4; controls: mean = 4.0; $t(24) = 5.49$ $p < .001$), while Chinese phrases and controls were seen as equally plausible (idioms: mean = 3.5; controls: mean = 3.4; $t(24) = 1.49$, $p = .15$). This suggests that plausibility was not driving the effects for 'unknown' items. If plausibility was affecting Chinese native speakers reading English phrases, we would expect to see a significant slowdown for controls, rather than simply a null effect. Similarly, the Chinese items are equally plausible in their idiom or control forms to naïve readers (English native speakers), hence the only way a difference can emerge is if some underlying knowledge of the idioms is being utilised, as in the case of the Chinese native speakers. We further explore the effect of plausibility in the models below.

We fitted separate models to compare the effects of familiarity, compositionality and plausibility. All continuous predictor variables were centred. We considered Chinese native speaker and English native speaker participants separately. In each model language and type were fixed effects and the interaction with each variable of interest was considered individually. Random intercepts for subject and item and by-subject random slopes were included for each fixed effect. Models were fitted for all word and phrase-level measures but only significant effects are described in detail here. (Full model outputs are provided in Appendix S2, tables S4–10, Supplementary Materials).

Familiarity

Subjective familiarity did not show significant effects for Chinese native speakers for Chinese idioms or English idioms. For English native speakers there was a marginal effect of familiarity on likelihood of skipping ($\beta = 0.29$, SE = 0.16, $z = 1.87$, $p = .06$). Closer inspection reveals that this reflects an interaction of familiarity and type for English idioms only (separate model for English phrases only, $z = -1.86$, $p = .06$). This pattern is repeated (although does not reach significance) for the later measures total dwell time and regression path duration. Hence for idioms, familiarity is facilitatory (more likely to skip, less likely to spend time re-reading the phrase or word). Conversely, controls of better known items are more likely to be read and re-read, presumably because the high familiarity generates a stronger expectation, the breaking of which is more problematic than for an idiom where the expected word is less strongly predicted. No significant effects were seen for Chinese items.

Compositionality

Compositionality showed no effects for Chinese native speakers for either set of phrases. This was also true of the compositionality ratings gathered from Chinese native speakers. English native speakers showed no effects of compositionality on any measure for English or Chinese items.

Plausibility

Plausibility showed no effect for Chinese native speakers reading English phrases, but was significant for Chinese phrases on early measures. For first fixation duration there was a significant interaction with phrase type ($\beta = 0.08$, SE = 0.04, $t = 1.95$, $p = .05$). This shows that more plausible phrases were read more quickly when the final word was correct, while for control phrases greater plausibility had an inhibitory effect. This trend is also seen in first pass reading time and total dwell time, although neither reaches significance. This means that for Chinese native speakers, who know the 'correct' completion, there is a clear difference in the effect of plausibility between the two variants. Crucially, when

CAMBRIDGE JOURNALS
http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81

Table 5. *Omnibus mixed effects model output for phrase-level reading patterns.*

	First pass reading time			Total reading time			Regression duration (word)			Regression duration (phrase)		
Fixed effects:	β	SE	t	β	SE	t	β	SE	t	β	SE	t
Intercept	6.62	0.17	39.37***	6.96	0.16	44.61***	6.27	0.10	65.36***	6.14	0.18	34.10***
Group: English	−0.55	0.08	−6.71***	−0.46	0.10	−4.46***	−0.28	0.13	−2.20*	−0.51	0.10	−4.96***
Language: English	−0.30	0.09	−3.53***	−0.28	0.07	−3.97***	−0.07	0.09	−0.82	−0.40	0.08	−4.95***
Type: Control	0.02	0.06	0.30	0.03	0.04	0.65	0.11	0.07	1.62	−0.05	0.04	−1.14
Group*Language	−0.17	0.08	−2.11*	−0.44	0.06	−7.85***	−0.45	0.11	−4.28***	−0.16	0.06	−2.64**
Group*Type	−0.13	0.08	−1.62	−0.02	0.06	−0.03	−0.13	0.09	−1.41	0.07	0.06	1.18
Language*Type	−0.07	0.08	−0.92	−0.01	0.05	−0.09	−0.07	0.09	−0.80	0.04	0.06	0.67
Group*Language*Type	0.31	0.11	2.91**	0.23	0.07	3.15**	0.29	0.13	2.28*	−0.02	0.08	−0.28
Control predictors:												
Word length (letters)	0.08	0.03	2.78**	0.09	0.03	3.66***	n/a	n/a	n/a	0.15	0.03	5.13***
Random effects	Variance			Variance			Variance			Variance		
Subject	0.038			0.090			0.117			0.087		
Item	0.034			0.029			0.026			0.044		
Subject | Language	0.006			0.005			0.027			0.006		
Subject | Type	0.003			0.007			0.010			0.002		
Residual	0.329			0.154			0.426			0.167		

Significance values are estimated by the R package lmerTest (version 2.0–11; Kuznetsova, Brockhoff & Christensen, 2014): *** $p < .001$, ** $p < .01$, * $p < .05$

CAMBRIDGE JOURNALS
http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81

reading Chinese phrases, English native speakers show the same pattern for both idioms and controls: as they have no underlying knowledge of the 'correct' idiom, plausibility plays an equal role for idioms and controls. In other words, *draw a snake and add . . .* can just as logically be completed with *hair* as it can *feet*, hence the effect is the same for either version. This shows that English native speakers did not consider the idioms or controls to be inherently more plausible (supporting the rating data). For English native speakers reading English phrases there was a significant interaction of plausibility and phrase type for skipping rate ($\beta = -0.56$, SE $= 0.25$, $z = -2.20$, $p <$ .05). Hence greater plausibility increased the likelihood of skipping in idioms, whereas for other measures it had a generally facilitatory but non-significant effect on both idioms and control items.

Proficiency

A final set of models were fitted to assess the contribution of English proficiency level for Chinese native speakers, considered in terms of three variables: vocabulary test score, self-rated ability and estimated usage. Each proficiency measure was assessed in turn for its overall effect, then for its interaction with language and phrase type. No measure of proficiency had an effect for the final word or whole phrase, or on regression durations. This suggests that the Chinese native speakers were generally well-matched in their English proficiency, and this may explain why we see no effects here: comparable studies that have found an effect of proficiency (e.g., Ueno, 2009) have done so with a deliberate high/low proficiency group manipulation.

Discussion

The results show complementary patterns for English native speaker and Chinese native speaker participants. Consistent with findings throughout the idiom literature, English native speakers show significant facilitation for the final words of a known phrase compared to a control phrase. The fact that the effect was most clearly evidenced in the likelihood of skipping (31% for idioms) suggests that this was highly automatic behaviour. As a result of this relatively high skipping rate, the early reading measures did not show much difference, but total reading time also showed a significant advantage. Chinese native speakers showed no effect for English idioms, which is again consistent with the previous literature on non-native speakers processing formulaic sequences in the L2. The Chinese items were not processed differentially by English native speakers on any measure, and crucially there was no difference in the effect of plausibility for the idioms vs. controls – this demonstrates that there is fundamentally no reason to expect the correct completion (e.g., *feet*) over the control completion (e.g., *hair*) unless the idiom is known. There was a consistent difference across duration measures for the Chinese native speakers, suggesting that there was some degree of crosslinguistic influence that provided a boost to lexical access for the items that were known in the L1. The effect was most clearly seen in the early measure first fixation duration, suggesting a degree of bottom-up facilitation through something akin to an interactive-activation framework (as suggested by Cutter, Drieghe & Liversedge, 2014 for their results on spaced compounds); it was also seen in total reading time, but not in phrase-level reading times or regression path measures. This in turn suggests that the lexical activation provided by the idiom is enough to facilitate the correct word, but not enough to overcome any inherent ambiguity in the non-compositional phrases. We will explore this dissociation further in Experiment 2.

One possible issue is that the idioms in the study were relatively long, and in particular the Chinese items were on average longer than the English items (Chinese items = 5.3 words; English items = 4.0 words, $t(50) = -4.55$, $p < .001$). However, in none of the analyses was the length of the prime a significant factor, i.e., a facilitative effect for the final word was seen whether the prime was relatively short (three words, e.g., *wine and meat (friends)*) or relatively long (six words, e.g., *beat the grass to scare the (snake)*). This suggests that the advantage seen for the Chinese native speakers was not necessarily strategic, although it is not possible to rule this out. Whether the result of strategic, active prediction or automatic lexical priming, we interpret the fact that we saw an effect for Chinese native speakers as evidence of L1 influence, even though the phrases were entirely novel in terms of form.

Experiment 2

In Experiment 2 we wanted to examine how participants read figurative and literal uses of the same idioms. In Siyanova-Chanturia et al. (2011) native speakers showed no difference in reading times for literal or figurative uses of ditropic idioms, whereas for non-native speakers figurative uses were read more slowly than literal uses. This difficulty understanding non-compositional phrases in the L2 may indicate that either the figurative meanings of idioms are unknown to non-native speakers, hence there is no direct entry to access, or that if the idioms are known, they are not accessed directly in the same way as for native speakers, and consideration of the figurative meaning only occurs after the literal meaning has been rejected. For translated items, if the idiom advantage observed in Experiment 1 is the result of activation of the underlying L1 idiom entry, we would expect figurative and literal uses of the translated Chinese idioms to be read comparably by Chinese native speakers, since activating the idiom will presumably also make the semantic meaning of the phrase available. More specifically, they

CAMBRIDGE JOURNALS
http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81

will be processed in the same way as English native speakers read English idioms. English native speakers should show a complementary pattern: difficulty reading the figurative uses of translated Chinese idioms compared to the entirely compositional literal uses.

Methodology

Participants

Twenty-one English native speakers and 21 Chinese native speakers took part in Experiment 2, all from a similar population as Experiment 1.

Materials

The English idioms used in Experiment 1 were augmented with stimuli from Siyanova-Chanturia et al. (2011) to give an initial set of 35 items. Chinese idioms were selected with the same selection criteria as for Experiment 1 (literal translation gave a grammatical English phrase with congruent word order, final word was a single word translation equivalent), with the additional stipulation that all idioms had to be literally plausible. To confirm this we included all English and Chinese idioms in a norming study where 24 English native speakers judged on a seven-point Likert scale how acceptable each was in a literal context. The 20 English and 20 Chinese idioms that were judged most plausible were retained (all received mean scores of greater than 3.5).

The idioms were placed into short contexts to bias either the figurative or literal meaning. These were included in a further norming study to assess how acceptable each was as an English sentence: 36 English native speakers judged their acceptability on a seven-point Likert scale. English items were rated as very acceptable in both figurative and literal contexts (figurative, mean = 6.3/7; literal, mean = 5.7/7). Chinese idioms were rated as being very acceptable in the literal contexts (mean = 5.6/7) and less acceptable in their figurative contexts (mean = 3.8/7), which is not surprising given that the idioms are all unknown to English native speakers. Familiarity of all items was verified in a separate norming test with 10 English native speakers. All idioms were then included in further norming studies with English native speakers to assess compositionality (n = 20; Chinese idioms were also assessed by Chinese native speakers, n = 12, in the original Chinese, as in Experiment 1). Table 6 shows example stimuli used in figurative and literal contexts.

Idioms were divided into two counterbalanced lists so that each participant saw 10 English idioms of each type (figurative/literal), 10 translated Chinese idioms of each type and 40 filler items. Within each list the idioms/controls were matched for number of words in the phrase, length and frequency of the final word, and literal plausibility of the idioms. The lexical coverage of all contexts was assessed using the Vocab Profile tool on the LexTutor website. All contexts had lexical coverage of greater than 96% at the K2 level (meaning that 96% of words were within the 2000 most frequent English word families) and greater than 99% coverage at the K5 level. In each item the idiom appeared toward the middle of the second line of a three-line block of text.

Procedure

All procedures were the same as in Experiment 1, however this time we took the whole phrase as the unit of analysis.[6] Because each analysis area was several words long, first fixation duration was discounted and first pass reading time was retained as our only early measure, with total reading time and total fixation count used as late measures. We also included regression path duration as an additional late measure to examine how participants used the preceding context to help understand each idiom.

Following the main experiment, participants were asked to provide subjective familiarity ratings for each idiom. English native speakers found English items highly familiar (mean = 6.4/7) and Chinese items unfamiliar (mean = 2.3/7). Chinese native speakers found Chinese items highly familiar (mean = 6.6/7) and English items less familiar (mean = 4.4/7). Chinese participants again completed a language background questionnaire and vocabulary test.

Analysis

No participants were removed from the analysis and the same data cleaning procedure as in Experiment 1 was applied. All trials where track loss occurred were removed. For native speakers this accounted for 1.9% of the data. For non-native speakers, in addition to the removal of trials where track loss occurred, any items containing unknown vocabulary items were removed, accounting for 5.3% of the non-native speaker data overall. English native speakers scored 92% on comprehension questions and non-native speakers scored 87%, suggesting that the task was again adequately completed by both groups. As with Experiment 1, duration measures were log-transformed to reduce skewing and for fixation count data a poisson regression was applied to the raw values. Table 7 shows a summary of results for all measures.

An omnibus model was fitted in which fixed effects of group (Chinese native speakers vs. English native speakers), language (Chinese phrases vs. English phrases) and phrase type (figurative vs. literal) and their interactions were computed. By-subject and by-item

[6] Separate analysis to compare final word reading for figurative vs. literal contexts showed no significant differences on any measures for either set of stimuli for English native speakers or Chinese native speakers.

CAMBRIDGE JOURNALS
http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81

Table 6. *Examples of ditropic English and translated Chinese idioms used in figurative and literal contexts.*

English idiom	A piece of cake – "easy"
Figurative context	One of my hobbies is doing little jobs around the house. I find most things I try are a piece of cake if you make sure you have the right tools before you start.
Literal context	Yesterday I was in the canteen at work and I was very hungry. I really wanted to get a piece of cake for my lunch but I was good and just had a sandwich.
Chinese idiom	Add oil and vinegar – "to embellish a story"
Figurative context	I have a friend who always exaggerates whenever he tells stories. The problem is he tends to add oil and vinegar so it's hard to know whether or not to believe what he says.
Literal context	I read a really simple recipe for a salad dressing. You just chop up some garlic and then add oil and vinegar then you put it in the fridge until you need to use it.

Table 7. *Summary of reading patterns for whole phrases for all measures, Chinese native speakers and English native speakers.*

	Chinese phrases		English phrases	
	Figurative	Literal	Figurative	Literal
Chinese native speakers				
First pass reading time	1350	1213	878	773
Total reading time	1985	1807	1242	1115
Total fixation count	7.8	7.3	4.9	4.5
Regression duration	1157	1033	644	576
English native speakers				
First pass reading time	739	681	394	400
Total reading time	1139	978	494	523
Total fixation count	5.4	4.7	2.6	2.6
Regression duration	644	585	308	316

Data are mean values in ms for duration measures and raw values for fixation counts.

random intercepts and by-subject slopes for language and phrase type were included in all models. The covariate idiom length (in words) was included in all models where log likelihood tests showed that this significantly improved the fit. Table 8 shows the omnibus results for all measures.

All measures showed significant main effects of group (English native speakers showed shorter reading times and fewer fixations, all ts > 4, all ps $< .001$), language (for all speakers English idioms were read more quickly than translated Chinese items, all ts > 2, all ps $< .05$), and importantly phrase type (literal phrases were read faster than figurative phrases, all ts > 2, all ps $< .05$). To further explore the data, separate models were fitted for Chinese native speakers and English native speakers (provided in Appendix S2, tables S11–12, Supplementary Materials).

Chinese native speakers show a significant main effect of type for all items (all ts > 2, all ps $< .05$) and no interactions between language and phrase type, suggesting that literal (compositional) uses were easier to understand than figurative uses for all phrases. In line with Siyanova-Chanturia et al. (2011) this was true for English (L2) idioms, but was also the case for translations of Chinese idioms. Therefore, despite the suggestion in Experiment 1 that known word combinations were being recognised/activated, this does not seem to translate into a straightforward understanding of the phrase-level meaning. For English native speakers reading English idioms, the results differ according to whether or not the idiom is English or Chinese in origin. For measures where there is a main effect for phrase type, this interacts significantly with language, hence the Chinese but not the English items show longer reading times for figurative phrases. Specifically, pairwise comparisons show that English phrases are read comparably whether they are

CAMBRIDGE JOURNALS
http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81

Table 8. *Omnibus mixed effects model output for all eye-tracking measures.*

	First pass reading time			Total reading time			Total fixation count			Regression duration		
Fixed effects:	β	SE	t	β	SE	t	β	SE	z	β	SE	t
Intercept	6.77	0.13	51.50***	7.28	0.14	53.70***	2.03	0.07	28.47***	6.63	0.15	44.40***
Group: English	−0.64	0.08	−8.21***	−0.57	0.09	−6.22***	−0.38	0.08	−4.73***	−0.69	0.10	−7.23***
Language: English	−0.38	0.07	−5.12***	−0.46	0.08	−6.07***	−0.46	0.08	−5.94***	−0.60	0.09	−6.82***
Type: Literal	−0.16	0.06	−2.78**	−0.12	0.04	−2.70**	−0.09	0.04	−2.33*	−0.15	0.05	−3.23***
Group*Language	−0.14	0.08	−1.70	−0.34	0.06	−5.49***	−0.27	0.07	−3.97***	−0.00	0.01	−0.08
Group*Type	0.06	0.08	0.71	−0.03	0.06	−0.55	−0.05	0.06	−0.91	−0.09	0.03	1.41
Language*Type	0.03	0.08	0.32	−0.02	0.06	−0.26	−0.02	−0.06	−0.31	−0.04	0.00	0.62
Group*Language*Type	0.09	0.11	0.81	0.18	0.08	2.24*	0.18	0.10	1.88	−0.03	0.00	0.33
Control predictors:												
Idiom length (words)	0.06	0.02	2.44*	0.04	0.02	1.86	n/a	n/a	n/a	0.07	0.02	2.70**
Random effects	Variance			Variance			Variance			Variance		
Subject	0.029			0.069			0.050			0.076		
Item	0.019			0.037			0.045			0.053		
Subject | Language	0.001			0.004			0.002			0.011		
Subject | Type	0.002			0.001			0.000			0.003		
Residual	0.316			0.165			n/a			0.185		

Significance values are estimated by the R package lmerTest (version 2.0–11; Kuznetsova, Brockhoff & Christensen, 2014): *** $p < .001$, ** $p < .01$, * $p < .05$

CAMBRIDGE JOURNALS
http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81

used figuratively or literally (all *p*s > .05), whereas there is a general slowdown for the figurative (non-compositional) uses of translated Chinese idioms. This is seen most clearly in the effect of type for Chinese phrases for total reading time (χ^2 (1, 822) = 13.39, $p < .001$) and total fixation count (χ^2 (1, 822) = 9.23, $p < .01$).

Familiarity, Compositionality, Plausibility

Separate models were fitted to assess the importance of these factors. Continuous predictor variables were centred. Chinese and English native speakers were considered separately, so models included language and phrase type as fixed effects and considered the interaction with each predictor variable in turn. Random intercepts for subject and item and by-subject random slopes for language and type were again included. Only significant results are discussed in detail below (all model outputs provided in Appendix S2, tables S13–19, Supplementary Materials).

Familiarity

For Chinese native speakers familiarity was not a significant factor in how Chinese or English items were read. Similarly, English native speakers showed no significant effects of familiarity for either set of items on any measures. Although this might seem surprising, the fact that all items in the study were deliberately chosen to be highly familiar may explain this (especially for the English items). In other words, items were either well known and were facilitated or were unknown and were not, with no 'sliding scale' of facilitation.

Compositionality

For Chinese native speakers compositionality played a role only in later measures. There was a significant interaction with phrase type for total reading time ($\beta = 0.12$, SE = 0.03, $t = 3.63$, $p < .001$) and total number of fixations ($\beta = 0.11$, SE = 0.03, $t = 3.34$, $p < .001$) and a significant three way interaction with phrase type and language for total reading time ($\beta = -0.20$, SE = 0.08, $t = 2.39$, $p < .05$). In both cases, greater compositionality was facilitatory for figurative and inhibitory for literal Chinese items, whereas in English the effect was facilitatory for literal items and negligible for figurative uses. For Chinese native speakers we also considered an alternative measure of compositionality, as judged by Chinese natives for the idioms read in the original Chinese characters. When these ratings were considered, greater compositionality was facilitatory for figurative uses for total reading time (interaction with phrase type: $\beta = 0.09$, SE = 0.04, $t = 2.24$, $p < .05$) and fixation count ($\beta = 0.10$, SE = 0.04, $t = 2.47$, $p < .05$) and showed no effect for literal items.

English native speakers showed significant interactions between type and compositionality and language and compositionality across all measures. This meant that for all items (Chinese and English phrase), more compositional items were actually read slower in the control condition, whilst the effect for figurative uses was negligible.

Plausibility

Literal plausibility (how acceptable each idiom would be if used in a literal context) showed a clear main effect for both groups for the Chinese items (for all measures except first pass reading for Chinese speakers, all *t*s > 2, all *p*s < .05). In all cases both figurative and literal uses were significantly facilitated by being more literally plausible, but there was no significant interaction between literal plausibility and phrase type. For Chinese speakers reading English idioms, both literal and figurative uses were also significantly facilitated by increased literal plausibility; for English native speakers there was facilitation for literal English phrases across all measures but not figurative phrases.

Proficiency

Models were fitted to assess the effect of vocabulary test scores, self-rated ability and usage scores for the Chinese native speakers. Usage was not significant, but both vocabulary test score and self-rated ability had a significant effect on all late measures (total reading time, regression path duration and fixation count). There was no interaction with language or phrase type, so higher proficiency led to faster reading across the board (which is expected), but participants were overall well-matched in their knowledge of the idioms. Increased proficiency did not therefore cause participants to read idioms from either language in a different way (more like native speakers), at least within the relatively homogenous cohort investigated here.

Discussion

Native English speakers performed as predicted. Idioms were read equally quickly in figurative and literal contexts, suggesting that, at least for the highly familiar idioms used here, there is no difference between a compositional analysis of the literal meaning and retrieval of the figurative meaning: both are available at around the same time. Chinese idioms, being unfamiliar to English speakers, were read significantly slower in figurative contexts, suggesting that their non-compositionality and the lack of a known figurative concept made them difficult to understand.

Chinese native speakers displayed the same pattern for both English and Chinese idioms: the literal versions of phrases were read more quickly than the figurative equivalents. This suggests that the overall meaning of the literal phrases could be understood with little difficulty,

CAMBRIDGE JOURNALS
http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81

whereas the non-compositional figurative uses were harder to integrate into the overall context. For the English stimuli this result is in line with comparable previous studies (e.g., Cieślicka, 2006; Siyanova-Chanturia et al., 2011) that show a clear advantage for literal over figurative meaning for non-native speakers. The fact that this pattern seems to hold for the translated Chinese stimuli raises some interesting questions. Despite the apparent priming of known lexical combinations in Experiment 1, the figurative meanings of the translated Chinese phrases were still difficult to understand in context, leading to longer overall reading times, more re-reading and more fixations. Therefore, although some degree of lexical activation seems to occur for the translated items, it may not be the case that the underlying conceptual entries associated with the idioms are automatically activated.

General Discussion

The combined results of Experiments 1 and 2 provide novel data on a previously underexplored aspect of the bilingual lexicon. Experiment 1 suggested that the recognition of the component words of idioms is an automatic process, even when the idiom is encountered in an unfamiliar translated form. This was seen specifically in the early reading behaviour for Chinese native speakers, where recognition of the 'correct' word was significantly and consistently faster than an unexpected control word. L1 influence must be important, as this is the only factor that renders the Chinese idioms 'known' for non-native speakers and not for English native speakers. Experiment 2, however, suggests that this recognition of form does not automatically lead to the activation of meaning: Chinese native speakers showed some difficulty interpreting the figurative phrases that were English idioms (as expected) and showed the same pattern for Chinese idioms. This was most clearly shown in total reading times, which reflect how easily the phrase can be integrated into the overall discourse context. This was also hinted at in Experiment 1, where phrase-level reading times and regression path durations were comparable for idioms and control phrases, suggesting that simply recognising the correct words did not prevent the Chinese native speakers from having to re-read the phrases to make any sense of them.

Based on these results, a conceptual route whereby idioms are represented in a language non-specific way seems unlikely; if this was the case, Chinese idioms should be understood relatively easily in translation. However, one important question relates to whether the figurative of literal meaning of an idiom is more salient, with the most salient in any given context being the one that is accessed first (Giora, 1997). As non-natives will almost always have encountered the component parts of idioms separately and in literal contexts more than in combination as an idiom, a literal, compositional reading is likely to be the default, and will remain the most salient interpretation until much higher levels of proficiency are reached (Cieślicka, 2006; Matlock & Heredia, 2002). For this reason we might expect a different set of results for participants of very advanced proficiency in terms of their reading of both English and Chinese phrases. As the participants in this study were all from the same cohort, this may explain the lack of any effect of proficiency level on the processing of the different phrase types.

Based on the advantage for the correct lexical forms seen in Experiment 1, the lexical-translation mechanism of idiom activation outlined in the introduction seems more plausible, but this is also not without its problems. If we assume that English stimuli are being quickly and automatically translated into Chinese and that this is triggering a known sequence, logically this should show some activation for the underlying concept. Thus, if a Chinese–English bilingual reading *draw a snake and add . . .* is quickly activating the Chinese equivalents and priming the character sequence 畫蛇添足, the conceptual meaning of this stored L1 form should be available alongside the final character, so making sense of the phrase in a figurative context in English should not be problematic. One explanation for the pattern of results is provided by more recent developments in idiom models, which suggest that idioms actually exist as multiple entries in the mental lexicon (Holsinger, 2013). In other words, they exist as distributed representations of single words with strong associative links, but also as canonical structures with set meanings. Thus the priming effect we see in activating the form of an idiom may be the result of lexical facilitation among the individual parts, whereas the representation of a whole form structure and its associated figurative meaning is likely to be affected by familiarity and (language specific) frequency of encounter for any given speaker. For native speakers, strong intralexical links and strong whole form representations exist to allow easy activation of both the form and meaning of the idiom. For Chinese native speakers, representations of whole forms are likely to be much weaker, both for L2 idioms and translations of L1 idioms, neither of which will have been regularly encountered in English. L2 idioms therefore do not show any lexical priming effects (Experiment 1) and are more difficult to process when used figuratively than literally because links with underlying concepts, if they exist, are not strong (Experiment 2). For the translated idioms, fast automatic translation may be sufficient to trigger associations through simple lexical priming/spreading activation, thereby facilitating formal recognition (Experiment 1), but the less salient, non-canonical presentation may not be sufficient to also trigger the whole form structure/meaning units (Experiment 2), or the novelty of encountering this form in English may work against its recognition.

CAMBRIDGE JOURNALS
http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81

Overall, it seems likely that idioms do retain some level of cohesion in translation. We interpret our findings in Experiment 1 as evidence that 'correct' completions were being primed, even though the idioms themselves were unknown in their translated forms. As demonstrated in Experiment 2, this activation did not extend to the overall meaning of the idioms, suggesting that the processes underlying recognition of form and access to meaning may not be the same, or that the 'compositional by default' approach for non-native speakers may negate any possible idiom advantage in the L2 until much higher levels of proficiency are reached. This study adds to previous work on the facilitative effect of congruence in formulaic language and provides suggestive evidence of crosslinguistic interaction at the multiword level, which adds a valuable new dimension to our understanding of the bilingual lexicon.

Supplementary Material

For supplementary material accompanying this paper, visit http://dx.doi.org/10.1017/S1366728915000103

References

Altarriba, J., Kroll, J., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory and Cognition, 24,* 477–492.

Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning: the effect of familiarity on children's repetition of four-word combinations. *Psychological Science, 19,* 241–248.

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

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R., Singmann, H., & Dai, B. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.0-7.

Bybee, J. (2006). From Usage to Grammar: The Mind's Response to Repetition. *Language, 82,* 711–733.

Bybee, J. (2008). Usage-based Grammar and Second Language Acquisition. In P. Robinson, & N. C. Ellis (eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*, pp. 216–236. New York: Routledge.

Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language, 27,* 668–683.

Caillies, S., & Butcher, K. (2007). Processing of Idiomatic Expressions: Evidence for a New Hybrid View. *Metaphor and Symbol, 22,* 79–108.

Carrol, G., & Conklin, K. (2014). Getting your wires crossed: Evidence for fast processing of L1 idioms in an L2. *Bilingualism: Language and Cognition, 17,* 784–797.

Charteris-Black, J. (2002). Second Language Figurative Proficiency: A Comparative Study of Malay and English. *Applied Linguistics, 23,* 104–133.

Chen, H., & Ng, M. (1989). Semantic facilitation and translation priming effects in Chinese–English bilinguals. *Memory and Cognition, 17,* 454–462.

Cieślicka, A. B. (2006). Literal salience in on-line processing of idiomatic expressions by second language learners. *Second Language Research, 22,* 115–144.

Cieślicka, A. B. (2013). Do nonnative language speakers chew the fat and spill the beans with different brain hemispheres? Investigating idiom decomposability with the divided visual field paradigm. *Journal of Psycholinguistic Research, 42,* 475–503.

Cieślicka, A. B., & Heredia, R. R. (2011). Hemispheric asymmetries in processing L1 and L2 idioms: Effects of salience and context. *Brain and Language, 116,* 136–150.

Cieślicka, A. B., Heredia, R. R., & Olivares, M. (2014). It's All in the Eyes: How Language Dominance, Salience, and Context Affect Eye Movements During Idiomatic Language Processing. In M. Pawlak & L. Aronin (eds.), *Essential Topics in Applied Linguistics and Multilingualism*, pp. 21–42. Switzerland: Springer International Publishing.

Clark, H. H. (1973). The language as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12,* 335–359.

Conklin, K., & Schmitt, N. (2008). Formulaic Sequences: Are They Processed More Quickly than Nonformulaic Language by Native and Nonnative Speakers? *Applied Linguistics, 29,* 72–89.

Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2014). Preview benefit in English spaced compounds. *Journal of Experimental Psychology: Learning, Memory and Cognition, 40,* 1778–1786.

De Groot, A., & Nas, G. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language, 30,* 90–123.

De Rosario-Martinez, H. (2013). Post hoc interaction analysis (phia). R package version 0.1-5.

Erman, B., & Warren, B. (2000). The idiom principle and the open-choice principle. *Text, 20,* 29–62.

Gibbs, R. W. (1991). Semantic Analyzability in children's understanding of idioms. *Journal of Speech and Hearing Research, 34,* 613–620.

Gibbs, R. W., Nayak, N. P., & Cutting, C. (1989). How to Kick the Bucket and Not Decompose: Analyzability and Idiom Processing. *Journal of Memory and Language, 28,* 576–593.

Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics, 8,* 183–206.

Google Translate. Available from: http://translate.google.co.uk/ Retrieved 10/06/2013.

Holsinger, E. (2013). Representing Idioms: Syntactic and Contextual Effects on Idiom Processing. *Language and Speech, 56,* 373–394.

CAMBRIDGE JOURNALS

http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81

Holsinger, E., & Kaiser, E. (2010). Effects of Context on Processing (non)-Compositional Expressions. 84th Annual Meeting of the Linguistics Society of America Baltimore, MD.

Inhoff, A. (1984). Two Stages of Word Processing during Eye Fixations in the Reading of Prose. *Journal of Verbal Learning and Verbal Behavior, 23,* 612–624.

Irujo, S. (1986). A piece of cake: learning and teaching idioms. *ELT Journal, 40,* 236–242.

Irujo, S. (1993). Steering clear: Avoidance in the production of idioms. *IRAL, 31,* 205–219.

Isobe, Y. (2011). Representation and Processing of Formulaic Sequences in L2 Mental Lexicon: How Do Japanese EFL Learners Process Multi-word Expressions? *JACET Kansai Journal, 13,* 38–49.

Jaeger, T. F. (2008). Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language, 59,* 434–446.

Jiang, N., & Nekrasova, T. (2007). The Processing of Formulaic Sequences by Second Language Speakers. *The Modern Language Journal, 91,* 433–445.

Kellerman, E. (1977). Towards a characterization of the strategy of transfer in second language learning. *Interlanguage Studies Bulletin, 2,* 58–145.

Kellerman, E. (1983). Now you see it, now you don't. In S. Gass, & L. Selinker (eds.), *Language transfer in language learning*, pp. 112–134. Rowley, MA: Newbury House.

Kellerman, E. (1986). An eye for an eye: crosslinguistic constraints on the development of the L2 lexicon. In E. Kellerman & M. Sharwood Smith (eds.), *Crosslinguistic Influence in Second Language Acquisition*, pp. 35–48. Oxford: Pergamon Press.

Kuznetsova, A., Brockhoff, P., & Christensen, R. (2014). lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-11.

Laufer, B. (2000). Avoidance of Idioms in Second Language: The Effect of L1–L2 Degree of Similarity. *Studia Linguistica, 54,* 186–196.

Libben, M., & Titone, D. (2008). The multidetermined nature of idiom processing. *Memory and Cognition, 36,* 1103–1121.

Lin, M., & Leonard, S. (2012). *Dictionary of 1000 Chinese Idioms (Revised Edition)*. New York: Hippocrene Books.

Liontas, J. (2001). That's all Greek to me! The comprehension and interpretation of modern Greek phrasal idioms. *The Reading Matrix, 1,* 1–32.

Liu, Y., Li, P., Shu, H., Zhang, Q., & Chen, L. (2010). Structure and meaning in Chinese: An ERP study of idioms. *Journal of Neurolinguistics, 23,* 615–630.

Matlock, T., & Heredia, R. R. (2002). Understanding phrasal verbs in monolinguals and bilinguals. In R. Heredia & J. Altarriba (eds.), *Bilingual Sentence Processing*, pp.251–274. Amsterdam: Elsevier.

McGlone, M., Glucksberg, S., & Cacciari, C. (1994). Semantic productivity and idiom comprehension. *Discourse Processes, 17,* 167–190.

Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher, 31,* 9–13.

On-line Chinese Tools Chinese/English Dictionary. Available from: http://www.mandarintools.com/worddict.html. Retrieved 10/06/2013.

Paterson, K., Liversedge, S., & Underwood, G. (1999). The influence of focus operators on syntactic processing of short relative clause sentences. *The Quarterly Journal of Experimental Psychology, 52a,* 717–737.

R Core Team (2014). R: A language and environment for statistical computing (version 3.1.2). R Foundation for Statistical Computing, Vienna, Austria.

Rommers, J., Dijkstra, T., & Bastiaansen, M. (2013). Context-dependent semantic processing in the human brain: evidence from idiom comprehension. *Journal of Cognitive Neuroscience, 25,* 762–76.

Schweigert, W. (1986). The comprehension of familiar and less familiar idioms. *Journal of Psycholinguistic Research, 15,* 33–45.

Schweigert, W. (1991). The Muddy Waters of Idiom Comprehension. *Journal of Psycholinguistic Research, 20,* 305–314.

Schweigert, W., & Moates, D. (1988). Familiar Idiom Comprehension. *Journal of Psycholinguistic Research, 17,* 281–296.

Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research, 27,* 251–272.

Smolka, E., Rabanus, S., & Rösler, F. (2007). Processing Verbs in German Idioms: Evidence Against the Configuration Hypothesis. *Metaphor and Symbol, 22,* 213–231.

Sprenger, S., Levelt, W., & Kempen, G. (2006). Lexical access during the production of idiomatic phrases. *Journal of Memory and Language, 54,* 161–184.

Staub, A., & Rayner, K. (2007). Eye movements and on-line comprehension processes. In M. G. Gaskell (ed.), *The Oxford Handbook of Psycholinguistics*, pp. 327–342. Oxford: Oxford University Press.

Swinney, D., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behaviour, 18,* 523–534.

Tabossi, P., Fanari, R., & Wolf, K. (2009). Why are idioms recognized fast? *Memory and Cognition, 37,* 529–40.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign language comprehension. *Proceedings of the National Academy of Science, USA, 104,* 12530–12535.

Titone, D., & Connine, C. (1999). On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics, 31,* 1655–1674.

Titone, D., Columbus, G., Whitford, V., Mercier, J., & Libben, M. (2015). Contrasting Bilingual and Monolingual Idiom Processing. In R. R. Heredia & A. B. Cieślicka (eds.), *Bilingual Figurative Language Processing*, pp. 171–207. Cambridge: Cambridge University Press.

Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing Advantages of Lexical Bundles: Evidence From Self-Paced Reading and Sentence Recall Tasks. *Language Learning, 61,* 569–613.

CAMBRIDGE JOURNALS

http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81

Ueno, T. (2009). An investigation of the relationship between the development of bilingual semantic organisation and interactive connectivity across languages. Unpublished PhD Thesis, Trinity College, Dublin.

Van Lancker Sidtis, D. (2012). Two-track mind: Formulaic and novel language support a dual-process model. In M. Faust (ed.), *The Handbook of the Neuropsychology of Language*, pp. 342–367. Chichester: Wiley-Blackwell.

Van Lancker, D., Canter, G. J., & Terbeek, D. (1981). Disambiguation of ditropic sentences: Acoustic and phonetic cues. *Journal of Speech and Hearing Research, 24,* 330–335.

Vocab Profile English – Compleat Lexical Tutor. Available at: http://www.lextutor.ca/vp/eng/.

Wang, X. (2007). Does Semantic Activation Spread Across Languages? An Experimental Study with Chinese–English Bilinguals. *Coyote Papers, 15,* 90–99.

Warren, H. (1994). *Oxford Learner's Dictionary of English Idioms*. Oxford, Oxford University Press.

Wolter, B., & Gyllstad, H. (2011). Collocational Links in the L2 Mental Lexicon and the Influence of L1 Intralexical Knowledge. *Applied Linguistics, 32,* 430–449.

Wolter, B., & Gyllstad, H. (2013). Frequency of Input and L2 Collocational Processing: A Comparison of Congruent and Incongruent Collocations. *Studies in Second Language Acquisition, 35,* 451–482.

Wolter, B., & Yamashita, J. (2014). Processing collocations in a second language: A case of first language activation? *Applied Psycholinguistics*, doi: 10.1017/S0142716414000113. Published online by Cambridge University Press, May 15, 2014.

Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge, Cambridge University Press.

Wray, A. (2012). What Do We (Think We) Know About Formulaic Language? An Evaluation of the Current State of Play. *Annual Review of Applied Linguistics, 32,* 231–254.

Wray, A., & Perkins, M. (2000). The functions of formulaic language: an integrated model. *Language and Communication, 20,* 1–28.

Wu, Y. J., Cristino, F., Leek, C., & Thierry, G. (2013). Non-selective lexical access in bilinguals is spontaneous and independent of input monitoring: evidence from eye tracking. *Cognition, 129,* 418–25.

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

Zhang, H., Yang, Y., Gu, J., & Ji, F. (2013). ERP correlates of compositionality in Chinese idiom comprehension. *Journal of Neurolinguistics, 26,* 89–112.

Zhang, T., Van Heuven, W. J., & Conklin, K. (2011). Fast automatic translation and morphological decomposition in Chinese–English bilinguals. *Psychological Science, 22,* 1237–42.

Zhou, S., Zhou, W., & Chen, X. (2004). Spatiotemporal Analysis of ERP During Chinese Idiom Comprehension. *Brain Topography, 17,* 27–37.

CAMBRIDGE JOURNALS

http://journals.cambridge.org Downloaded: 21 Apr 2015 IP address: 128.243.139.81