Title	Naturalistic Reading in the L2 and the Impact of Word Frequency and Cross-Linguistic Similarity
Author(s)	DAVID, Allen; KATHY, Conklin
Citation	Journal of the Ochanomizu University English Society
Issue Date	2017-03
URL	http://hdl.handle.net/10083/60790
Rights	
Resource Type	Departmental Bulletin Paper
Resource Version	publisher
Additional Information	

Naturalistic Reading in the L2 and the Impact of Word Frequency and Cross-Linguistic Similarity

David Allen and Kathy Conklin

Abstract

While psycholinguistic studies of first language (L1) reading have identified multiple factors that predict the speed of lexical access, there are few studies investigating whether such factors influence second language (L2) reading. For usage-based models of acquisition and processing, two lexical factors that are believed to be crucial in L2 reading are word frequency and cross-linguistic similarity. No previous studies, however, have looked at these factors during naturalistic reading tasks when readers' L1 and L2 differ in script. In this study, we monitored the eye movements of Japanese speakers of English while they read a short story. We used linear mixed effects modelling to investigate the role of word frequency and cross-linguistic similarity, as well as other factors such as language proficiency, on L2 lexical access. Word frequency was a strong predictor of word reading speed. A cross-linguistic measure of phonological similarity was not significant, indicating that even if lexical representations in the L1 were activated during L2 reading, this activation did not influence reading speed. The findings are discussed in terms of a localist connectionist model of word recognition.

Keywords: Bilingual lexicon, Japanese-English cognates, cross-linguistic activation, eye-tracking, word frequency

Introduction

A fundamental question in second language research is how similar second language (L2) processing is to first language (L1) processing. According to a usage-based theory, in both an L1 and L2, frequency of input underpins the development of linguistic knowledge and its organization, as well as having a profound effect on processing (e.g., Bybee & Hopper, 2001; Ellis, 2002). Such a theory predicts that words that are encountered more frequently, be that in an L1 or L2, will be processed more quickly. While in this way L1 and L2 processing are believed to be fundamentally the same, speakers of two languages confront a challenge that monolinguals do not; namely when using the L2, the L1 is often activated, which can influence processing (Dijkstra, 2007). In short, because language learners already have an established language, its influence is often evidenced in an L2 or L3 (e.g., Odlin, 1989; Ringbom, 1987).

The present study investigates the impact of word frequency and cross-linguistic overlap (via the use of cognates) during L2 reading. While previous studies have tended to focus on the influences of these factors in single-word tasks such as lexical decision, we focus on their influence while participants read a story in the L2. Moreover, we believe this study is the first to investigate such effects during naturalistic reading when the bilinguals' languages differ in script (i.e., Japanese and English)[1]. How lexical and cross-linguistic factors impact processing during such natural reading is important as such reading tasks are commonly used both inside and outside of classrooms, and in language assessment situations. Our study considers the findings in terms of a connectionist model of bilingual lexical processing.

Word frequency

The word frequency effect (FE) in psycholinguistic studies of lexical processing refers to the robust observation that words that are higher in frequency are recognized more quickly than otherwise matched lower frequency ones (Rayner, 1998). It is widely assumed that the FE reflects implicit learning (Ellis, 2002), where lexical representations are strengthened each time a word is seen; that is, the connections between its orthographic form, phonological form and its meaning all become strengthened as a function of exposure. This assumption is a fundamental one within emergent and usage-based theories of second language acquisition, and applies to single words, phrases (Ellis, Simpson-Vlach & Maynard, 2008; Tremblay, 2011) and constructions (Ellis, 1996, 2002).

Previous studies have demonstrated that word frequency is a crucial factor in lexical access during L2 word recognition (e.g., Duyck et al., 2008). The few L2 studies that have focused on lexical access in more naturalistic reading tasks have also observed the FE (Balling, 2013; Cop et al., 2015; Whitford & Titone, 2012). Moreover, two of these studies showed that the FE is greater when reading in the L2 than in the L1 (Cop et al., 2015; Whitford & Titone, 2012). Based on these studies, then, we expect to observe a frequency during longer reading tasks in the L2.

In localist connectionist models (e.g., McClelland & Rumelhart, 1981), the FE is accounted for by manipulating the thresholds required for activation of a particular lexical representation. High frequency words have lower thresholds for activation such that they become activated more quickly than low frequency words which have higher thresholds. The Bilingual Interactive Activation Plus model (BIA+; Dijkstra & van Heuven, 2002), which was based on McClelland and Rumelhart's (1981) original IA model for monolingual word recognition and is the currently dominant model of bilingual lexical processing, postulates this same mechanism for the FE in two languages. Within-language frequency effects for both L1 and L2 are thus accounted for via a manipulation of thresholds for activation: High frequency L2 words will have lower thresholds than low frequency L2 words.

---

[1] In this study we follow the definition of 'bilinguals' that is typically used in psycholinguistics (e.g., Dijkstra, 2007), that is, bilinguals are people who know two languages, regardless of the level of proficiency in those languages.

Cross-linguistic influences

When investigating cross-linguistic influence at the lexical level, the most important category of words is cognates. Cognates share some degree of form (phonology and / or orthography) and meaning (semantics) across languages. For example, the words *tomato* in English, *tomaat* in Dutch and トマト/tomato/ in Japanese all share meaning and some degree of form, making them cognate according to the prevalent definition in psycholinguistics (Dijkstra, 2007). For languages that share a script, such as Dutch and English, cognates have overlap in both orthography and phonology, while in languages that differ in script, such as Japanese and English, only phonological form is shared.

In same script languages, cognates can be either identical or similar in orthography (e.g., *metro-metro* and *tomaat-tomato* in English and Dutch). Cognate phonology, on the other hand, is rarely identical due to the fine-grained differences in phonological features across languages. Therefore, phonology can vary from very similar to somewhat similar for cognates. Because of this gradation in phonological similarity, researchers have begun to use continuous measures of formal overlap to define more precisely what it means to be ‘cognate’, rather than relying on a simple dichotomy of cognate/noncognate (e.g., Allen & Conklin, 2013, 2014; Dijkstra et al., 2010).

Research in psychology has shown cognates are processed more quickly and accurately than matched controls in reading tasks that present words in isolation, such as lexical decision (e.g., de Groot & Nas, 1991; Dijkstra, Grainger, & van Heuven, 1999). The robustness of this ‘cognate facilitation’ effect in L2 word recognition has been observed with a number of L1s and L2s (e.g., Spanish-Catalan, English-German, French-English, Spanish-Basque, Danish-English) and even when the L2 and L1 do not share a script, such as Japanese and English, or Korean and English (e.g., Allen & Conklin, 2013; Kim & Davis, 2003). Moreover, the degree of formal overlap has also been found to influence the magnitude of the facilitation, such that when words are more similar (ラジオ /rajio/ *radio* vs. テレビ /terebi/ *television*) they are processed more quickly (Allen & Conklin, 2013; Dijkstra et al., 2010; Van Assche et al., 2009; 2011).

The faster processing for L2 cognate words is hypothesised to be due to activation of L1 word representations when reading in the L2. Thus, the prevailing view of bilingual processing is that lexical access is inherently ‘non-selective’, as words in both languages are activated in the mental lexicon. The activation of shared formal and semantic features in the L1 and L2 boosts activation of the target word leading to faster responses by the reader. Connectionist models of bilingual word recognition (e.g., BIA+) have been reasonably successful in showing how this interaction between languages in the bilingual lexicon allows cognate facilitation to emerge (Dijkstra & Van Heuven, 2002).

The cognate facilitation effect has also been observed when targets are embedded in sentence contexts suggesting that the availability of context does not inhibit the activation of multiple languages (e.g., Bultena, Dijkstra, & Van Hell, 2013; Van Assche et al., 2011). On the other hand, some studies have found that cognate effects disappear or are reduced when cognates are in more predictable in context (Libben & Titone, 2009; Schwartz & Kroll, 2006; Van Hell & de Groot, 2008).

The processing advantage for cognates over noncognates in more naturalistic reading tasks has only been looked at in two recent studies (Balling, 2013; Cop et al., 2016). Balling (2013) had highly proficient Danish-English bilinguals read newspaper articles at their own pace while their eye-movements were recorded. A significant advantage for cognates was observed in the total reading time measure, suggesting that readers made fewer regressions to cognate targets and that fixations were shorter overall than for noncognate controls. More recently, Cop et al. (2016) showed that Dutch-English bilinguals processed cognates more quickly than noncognates when reading a novel in English. As in previous studies, greater orthographic overlap resulted in reduced fixation times. These studies suggest that the cognate facilitation effect holds for naturalistic reading tasks. However, both of these studies focused on languages that share script. An important question remains about whether the cognate advantage persists in naturalistic reading with languages that differ in script.

Because much of the cognate advantage found in shared script languages is thought to derive from orthographic similarity, and the influence of phonological similarity is considered secondary (e.g., Cop et al., 2016; Dijkstra et al., 2010), the cognate advantage may not be observable in naturalistic reading tasks in different script languages. The best way to investigate whether phonology is an important contributor to the effect is to explore whether facilitation remains when orthography is eliminated; Japanese-English cognates, which do not share a script, will allow us to address this question.

Importantly, Japanese-English cognates are ubiquitous (Shibatani, 1990) and the degree of phonological similarity for them has been shown to influence L2 (English) processing in tasks with words presented in isolation (Allen & Conklin, 2013; Dijkstra et al., 2010; Hoshino & Kroll, 2008). The present study asks whether this phonological facilitation remains in tasks that involve presentation of targets within a rich semantic context, such as a story. Moreover, we investigate whether and to what extent the FE facilitates lexical access during naturalistic reading in the L2.

Method

Participants

Eleven Japanese-English bilinguals (4 male, 7 female; mean age=26 years, SD=6 years) participated in the experiment. All participants had normal or corrected-to-normal vision. All participants were Japanese-English bilinguals who were enrolled in programs at a UK University. The participants were paid 7.50 GBP for their participation and completed informed consent forms prior to the experiment as per the ethics regulations at the institution. Following the eye-tracking experiment, participants completed a language history and experience questionnaire. Self-ratings (on a scale of 1-10, with 10 being the highest) for reading, writing, speaking and listening proficiency in both English and Japanese were collected and an average proficiency across these skills was calculated. This averaged measure is presented in Table 1 below. Additionally, age-of-acquisition (AoA) and length of stay information were collected to establish language history (Table 1).

Participants were proficient in both English (M=7.6, SD=2.2) and Japanese (M=8.8, SD=1.6). All participants began learning Japanese from birth (0 years) while English in all but two cases (participants 9

Table 1: Japanese/English proficiency, age-of-acquisition and length of stay information for participants

High L2 Proficiency

Participant Number	Age	English Proficiency	English AoA	Length of Stay in UK	Japanese Proficiency	Japanese AoA
7	22	10.0	6-10yr	10yr+	8	0yr
8	20	9.5	6-10yr	10yr+	8.75	0yr
9	20	9.8	0yr	10yr+	4.75	0yr
10	19	10.0	0yr	1-2yr*	9.5	0yr
12	23	9.0	6-10yr	10yr+	9.25	0yr
M (SD)	20.8 (1.6)	9.7 (0.4)			8.1 (1.9)	

Low L2 Proficiency

Participant Number	Age	English Proficiency	English AoA	Length of Stay in UK	Japanese Proficiency	Japanese AoA
2	34	6.0	6-10yr	<1yr	9.25	0yr
3	26	6.5	11-15yr	3-4yr	9.75	0yr
4	30	7.3	11-15yr	4-5yr	10	0yr
5	34	5.3	11-15yr	<1yr	10	0yr
6	31	7.0	11-15yr	<1yr	10	0yr
11	22	3.8	11-15yr	<1yr	7.5	0yr
M (SD)	29.5 (4.7)	6.0 (1.3)			9.4 (1.0)	

* Participant 10 had only stayed in the UK for 1-2 years but had an English-speaking parent who always used English with her.

and 10) was acquired later in life (starting between 6 and 15 years). However, proficiency varied, meaning that five participants rated themselves as more proficient in English or roughly equally proficient in both languages; in other words, five of the participants appeared to be balanced bilinguals, while the other six were clearly unbalanced, being more dominant in Japanese. This was reflected in the length of stay in the UK, with the higher English proficiency group tending to have resided in the UK for longer (over 10 years). Based on this information, the participants were grouped into a high and a low L2 proficiency group, of 5 and 6 participants, respectively. The two groups differed significantly in average English proficiency (High=9.7 (0.4); Low=6.0 (1.3); t=6.08, df=9, p<.001) but not in average Japanese proficiency (High=8.1 (1.9); Low=9.4 (1.0); t=1.52, df=9, p=0.16).

To further investigate the differences between the two proficiency groups, the on-line measures from the eye-tracking data, such as number of fixations and average saccade length, were compared using independent samples *t*-tests (Table 2). For the low proficiency group, the number of fixations and the number of saccades were significantly greater, and the average fixation duration and average reading time were both significantly longer, all indicating increased difficulty in processing the texts relative to the high proficiency group. The average saccade amplitude was significantly shorter for the low proficiency group suggesting lower reading fluency compared to the higher proficiency group.

Materials: Text and items

A fictional text (1120 words) was written by the researcher for the purposes of the experiment. The

Table 2: Measures of reading performance for the two proficiency groups

	High Proficiency (*n*=5)	Low Proficiency (*n*=6)	*t*-value	*p*-value
Number of fixations	72.2 (28.5)	96.8 (20.7)	6.163	<.001
Average fixation duration	223.3 (17.2)	229.7 (19.9)	2.085	<.05
Number of saccades	81.8 (22.9)	116.9 (37.4)	7.126	<.001
Average Saccade amplitude	4.8 (0.7)	4.2 (0.7)	-5.412	<.001
Average reading time	21010ms (6269ms) 21 mins (6 mins)	30762ms (10306ms) 31 mins (10 mins)	7.201	<.001

*Standard deviations in parentheses

content of the text, as well as the vocabulary, were selected to be accessible to all participants. To confirm the frequency of words in the text, it was analyzed using the online application LexTutor (http://www.lextutor.ca/cgi-bin/vp/comp/output.pl) and the BNC 1-20,000 wordlists. This showed that 83.2% of words fell into the top 1000-frequency band, 92.7% into the top 3000-, and 96.8% into the top 5000-frequency bands. Thirty-two words fell into bands below the 5000-word level.

The text, which is a simple narrative story, was written to include a range of unequivocal Japanese-English cognates (i.e., English words that are always translated as cognates in Japanese) as well as similarly unequivocal noncognates. While texts such as newspaper articles may be more appropriate for an authentic reading task, it proved difficult to find an existing English text that contained a sufficient number of items that could be reliably confirmed as Japanese-English cognates in context.

Fifty-two items made up the target items in the study (26 cognates and 26 noncognates; see Appendix 1). To confirm the cognate status of the words in the text, five highly proficient Japanese-English bilinguals, all of whom had done translation work in the past, performed a translation task. In the first part of the task, the bilinguals identified items that could *only* be translated as cognates in Japanese and those that did not have a recognizable cognate translation and therefore were *only* translatable into Japanese as noncognates. The identified items were selected as possible cognate and noncognate items, respectively. Then, the same five bilinguals translated each item from the new list into Japanese. Participants provided the most appropriate translation for each item in the context of the text. Cognates and noncognates were always translated as such. For example, *penguin* was identified as a cognate in Japanese (ペンギン/pengin/) that could not be translated using a noncognate alternative translation, and all bilinguals translated *penguin* using this cognate Japanese translation; on the other hand, *zoo* was identified as a noncognate in Japanese (動物園/doubutsuen/) that did not have a recognisable cognate translation (i.e., ズー/zuu/*) and was always translated using a noncognate translation.

These 52 items were selected as regions of interest (ROIs). All were content words and no items were at the beginning or end of lines or sentences. Target items appeared more than once in the text. However, analyses were only conducted on the first encounter of a target word, as repetition effects could influence subsequent encounters (e.g., Rayner et al., 1995). Target items were of various grammatical classes though the majority were nouns (nouns=25 items, noun/verbs=18, adjectives=5, noun/adjective=3, verb/adjective=1).

Apparatus and procedure

The experiment was conducted using an SMI Eye-Link 1 head-worn eye-tracker and Experiment Builder software (SR Research Ltd., 250Hz). The screen used for presenting materials had a resolution of 1280 x 1024 pixels. The texts were presented in 14-point Courier New, black text on white background and the lines were double-spaced. Each paragraph was presented on a separate page and participants pressed the spacebar on the keyboard to move to the next page.

Participants were fitted with the eye-tracker and calibration was performed using a 9-point grid. Instructions were presented orally and on-screen. Participants were told to read the text and answer seven comprehension questions that would follow, thereby focusing the participants on reading for comprehension. Seven statements were provided that targeted details (e.g., The weather was fine) and required a True/False response. A practice task was performed first, which served as a model for the main task. A second calibration was performed prior to the main reading task.

Analyses

An initial analysis investigated general measures of participants' reading performance for all words in the text in order to confirm group differences in L2 reading ability. Following this, the main analysis investigated which factors influenced L2 reading performance for the target lexical items.

Dependent measures

One early measure (Gaze Duration, henceforth GD) and one late measure (Total Reading Time, TRT) were used as the response variables. GD is the sum of all durations of fixations on target from the first fixation to the time of leaving the word. TRT is the total time spent fixated on the target word including regressions and refixations. It provides the best indicator of overall difficulty of the target words as it includes regressions to it following initial processing. Both GD and TRT measures were log-transformed to improve linearity and reduce random variance (base=*e;* Baayen, 2008).

Predictors of reading performance for individual items

The primary variables of interest were word frequency and cross-linguistic phonological similarity. Interactions between L2 proficiency and the other variables were included in the analysis in case the high/low proficiency groups exhibited different tendencies in reading processes.

Word frequency was taken from the BNC and log-transformed (base=*e*). The cross-linguistic phonological similarity measure was derived from bilinguals' ratings on a scale of 1 to 5 (1 = completely different, 5 = identical) of how phonologically similar translation pairs, such as *radio* – ラジオ/*rajio*/, were perceived to be (Allen & Conklin, 2014). Phonological similarity is a continuous measure of cognateness that provides more explanatory power than a binary measure (Allen & Conklin, 2013, 2014). Most of the cross-linguistic similarity ratings were taken from a database of Japanese-English translation equivalents (Allen & Conklin, 2014). However, a number of items did not have ratings for phonological similarity and so an additional rating study was conducted to collect these missing data. Eleven undergraduates at a Japanese university (5 female, 6 male; all undergraduate university students; mean age=21 years, SD=3 years) completed similarity-rating tasks, which were identical to those reported in Allen & Conklin (2014). Using the ratings, it was easy to distinguish cognates from noncognates (cognates M=3.4, SD=0.5; noncognate M=1.0, SD=0.1; $t = 32.04$, $df = 50$, $p<.001$), such that any item rated above 1.5 is cognate.

L2 proficiency was a two-level factorial variable (high vs. low) based on participants' self-rating of English proficiency. Additional lexical variables that previous research has shown to be important predictors of reading were included to control for their influence on reading times: Word length (e.g., Rayner, 1998), morphological complexity (e.g., Balling, 2013), context predictability (McDonald & Shillcock, 2003) and number of senses (Hino et al., 2002). Word length was measured by counting the number of letters in each word. Morphological complexity was measured by counting the number of morphemes in each word. For example, the English target words *curtains* and *purchased* both contained two morphemes (*curtain-s, purchase-d*). A context predictability measure was added which used the frequencies of word sequences from the BNC. This is based on the idea that the transitional probability of words, that is, how likely one word is to follow another, affects processing of upcoming words (McDonald & Shillcock, 2003). To calculate this measure, the frequency of the trigram in which the item appears as the last word was divided by the frequency of the bigram that precedes the target word (for example, for the target item *bed*, the frequency of the trigram *lying in bed* was divided by the frequency of the bigram *lying in*). Because zeros are common in the bi/trigram frequencies, a +1 transformation was applied to all frequencies in order to correct for these zero frequencies. The total number of English senses for words as listed online on the WordNet database (Princeton, 1990) was used to assess the impact of polysemy.

Finally, the numerical position of the word in the text (where the first word is 1) was included to assess the possible build up of contextual information that dictates overall reading speed. This factor also acts as an indicator of any reading/task fatigue, which would be demonstrated by reading times becoming slower as the text progresses.

Linear mixed effects modelling

To explore the contribution of the various factors, linear mixed effects modelling (Baayen, Davidson, & Bates, 2008) was conducted with R version 2.11.1 (R Core Development Team, 2010) and the R package lme4 version 0.999375-37 (Bates & Maechler, 2010). The above predictors and interactions were considered in the model and random intercepts of items and participants were included, as well as random slopes for all predictors and interactions (see Barr, Levy, Scheepers & Tily, 2013, for more details on the necessity of including a maximal random effects structure in linear mixed effects models).

To deal with the issue of natural correlations between predictor variables, such as frequency and length, a correlation analysis was performed. When two or more predictors were significantly correlated, we removed this collinearity through a process of residualization (see Baayen, 2008). The collinearity was removed by fitting a linear model in which one variable (e.g., frequency) became the response and its correlated variables (e.g., word length, number of senses) became the predictor variables. This procedure was repeated for all correlated variables so that each variable was the response variable in a model with all of its correlated predictors. The residuals of these models were used as the predictor variables in the final analyses. The resulting residuals were significantly correlated with their related variables (*p*<.01).

A backward simplification procedure was conducted, such that all main effects and interactions were included in the initial model and non-significant effects/interactions were removed step-by-step. Interactions were removed prior to main effects, and each time a term was removed a log-likelihood ratio test was performed to show whether this removal significantly affected the predictive capability of the model. If the removal lead to a significantly less explanatory model (*p*<.05), the term was retained; otherwise, it was removed.

Results

Data cleaning

Only fixations that fell into the interest areas defined around the target words were considered for analysis. Interest areas were defined using an auto-segment function in the DataViewer software. Single fixations that were shorter than 100ms were excluded, as likely reflecting oculomotor programming, while those over 800ms were removed, as likely being due to blinks or momentary track loss (Rayner, 1998). Fixations due to blinks and track loss but which had durations of less than 800ms were also identified and removed. To reduce any additional within-subject random variance, data points on both continuous dependent measures (GD, TRT) that were +/-2 SDs of each participant's mean for that measure were removed. This lead to 6.2% of data being removed. All participants, except for one answered the seven multiple-choice questions correctly (one participant answered one question incorrectly) immediately after finishing. This showed that all participants were reading the text for meaning comprehension and consequently all data could be used for statistical analyses.

Target items analysis

The final models for GD and TRT are presented below in Tables 3 and 4, respectively. The estimated coefficients of the fixed effects, the standard error, *t*-values and the *p*-values obtained from *t*-tests are presented. The standard deviation, and variance for random effects of participants and items (intercepts) as well as the slopes for the random variables for each significant predictor in the final model are shown in each table.

Log-transformed English frequency was highly significant in both GD and TRT models (*p*<.001), such that more frequent words were fixated on for shorter durations. The interaction between frequency and L2 proficiency was not significant and was thus removed during model simplification. Separate analyses conducted on the data for the high and low proficiency groups confirmed that the FE was highly significant for both GD and TRT for both groups (*p*<.001).

There was no significant effect of phonological similarity on GD or TRT, indicating no effect

Table 3: Final model for log-transformed gaze durations (GD) for targets

Random effects			
Groups	Name	Variance	SD
Item	*(Intercept)*	0.001	0.038
	Proficiency	0.002	0.043
Subject	*(Intercept)*	0.000	0.000
	Number of Senses	0.000	0.000
	Word frequency	0.000	0.000
	Word length	0.000	0.020
	Context predictability	0.005	0.074
	Position in text	0.000	0.000
Residual		0.170	0.412

Fixed effects				
	Estimate	Std. Error	*t*-value	*p*-value
(Intercept)	5.444	0.072	75.64	<0.001
L2 proficiency	0.247	0.076	3.23	0.013
Word frequency	-0.089	0.018	-4.93	<0.001
Word length	0.110	0.017	6.34	<0.001
Number of Senses	-0.013	0.006	-2.19	0.029
Context predictability	-0.150	0.069	-2.16	0.031
Position in text	0.000	0.000	2.30	0.022

Table 4: Final model for log-transformed total reading time (TRT) for targets

Random effects			
Groups	Name	Variance	SD
Item	*(Intercept)*	0.001	0.027
	L2 proficiency	0.013	0.115
Subject	*(Intercept)*	0.093	0.305
	Word frequency	0.000	0.011
	Word length	0.001	0.036
Residual		0.192	0.438

Fixed effects				
	Estimate	Std. Error	*t*-value	*p*-value
(Intercept)	5.691	0.059	96.21	<0.001
Word frequency	-0.081	0.018	-4.61	<0.001
L2 proficiency	0.279	0.081	3.45	0.006
Word length	0.115	0.019	6.00	<0.001

of cognate status on L2 reading with Japanese-English bilinguals. Similarly, the interaction between phonological similarity and L2 proficiency was not significant. Both the interaction and fixed effect terms were removed during model simplification. L2 proficiency as a fixed effect was significant in both GD and TRT final models ($p<.05$, $p<.01$, respectively), such that higher L2 proficiency led to reduced fixation times.

Word length was also highly significant in both GD and TRT models ($p<.001$), showing that longer words had longer fixation times. Both frequency and word length effects on reading times are well documented (e.g., Rayner et al., 1996), and establish that the current bilinguals were utilising L2 lexical properties in reading in a similar fashion to monolinguals.

The number of English senses was significant in the GD model only ($p<.05$), such that more-polysemous words were read more quickly than less-polysemous words. Context predictability was significant for GD only ($p<.05$) showing that target items that had higher probability of occurring with the previous two words were fixated on for shorter durations.

The number of morphemes was not significant and was thus removed from the model. The position of the word in the text was significant in the GD model ($p<.05$), showing a general slowing of word reading times as participants progressed through the text. This may indicate a fatigue effect, or alternatively, a slowing due to integration of more information as a story progresses.

Discussion

The present study found a strong FE during naturalistic L2 reading, where individual lexical items were processed in context. This supports the view that speed of lexical access during natural reading is strongly determined by lexical frequency and is in line with other studies using shared script languages (Balling, 2013; Cop et al., 2015; Whitford & Titone, 2012). These findings provide support for usage-based models of L2 acquisition and processing. They suggest that L2 learners utilize similar cognitive mechanisms for lexical processing as in the L1. This means that in the L2, form-meaning mappings in the lexicon become stronger through exposure, with repeated input leading to faster lexical access during reading tasks. This type of implicit learning leads to the FEs observed in naturalistic L2 reading.

In contrast to word frequency, no effect of phonological similarity was found on reading times. Thus, the cross-linguistic phonological similarity between Japanese and English cognate words, which has been shown to influence lexical processing in both single-word production (Hoshino & Kroll, 2008) and comprehension tasks (Allen & Conklin, 2013; Dijkstra et al., 2010), was inconsequential in this type of reading task. This finding contrasts with two recent studies that have observed cognate facilitation in naturalistic reading tasks with shared-script languages (Balling, 2013; Cop et al., 2016). In the following, we present an interpretation of this finding in terms of the BIA+ model for word recognition and we consider two alternative explanations for the lack of cross-linguistic similarity effects: The effect of context and boosted L2 semantics and the lack of shared orthography.

The findings can be interpreted in terms of the BIA+. Due to the non-selective nature of lexical activation in the bilingual lexicon (Dijkstra, 2007), when a word shares formal features across languages, activation of these features in both languages boosts the degree of activation of lexical representations in both languages. Thus, regardless of the type of reading task, bottom-up activation should occur across languages for cognates. Although the degree of phonological similarity did not measurably impact processing of words in the present study, the BIA+ model assumes that cross-linguistic activation *did* occur during lexical activation. However, during reading top-down global activation of the L2 may have attenuated the activation of lexical candidates in the L1, leading to a lack of observable advantage for cognate words. Moreover, if the greater semantic context provided by the story biases readers to L2 interpretations, top-down semantic activation of L2-specific word meanings may reduce the influence of bottom-up cross-linguistic phonological similarity. In other words, the L2 context may reduce the influence of the L1, manifesting in what appears to be more 'language-selective' processing. While these results do not necessarily indicate that more natural reading is language-selective, they do support the notion that rich language-specific contexts can lead to effects that resemble language-selective processing.

Importantly, two recent studies did find a cognate advantage in a very similar task (Balling, 2013; Cop et al., 2016). This makes it unlikely that strong top-down effects from rich language-specific contexts in naturalistic reading tasks mitigate bottom-up cross-linguistic activation. However, an important difference between the present study and these two previous studies is that Japanese and English do not share orthography whereas Danish, Dutch and English do. The BIA+ can account for the different pattern of results for shared- and different-script languages. If both orthography and phonology are shared, this should lead to a greater amount of cross-linguistic activation than if languages only share one of these. The current findings seem to support the view that there is less cross-linguistic activation when languages only share phonology. Thus for languages that do not share script, lexical access during naturalistic L2 reading appears to be minimally influenced by cross-linguistic phonological similarity. Having both orthographic and phonological similarity may be needed for cognates to be sufficiently activated for an advantage to become apparent.

Conclusions and limitations

The present study, we believe, is the first to investigate L2 natural reading by different-script bilinguals. We observed a strong FE, providing clear evidence for usage based models that hold that reading in the L2 is similar to reading in L1, even when bilinguals' first and second languages differ in script. Furthermore, we show that the ubiquitous cognate effect is not observed in naturalistic reading for different-script bilinguals, which could be because of the different scripts and/or because a rich semantic context helps mitigate the influence of the L1.

A limitation of the present study is that the interpretation of the findings is constrained by the size of the study. Eleven participants took part and 52 items were analyzed for each participant. The null effect of cross-linguistic similarity may, therefore, also be due to the limited sample size. Whereas observed frequency effects are typically large across reading studies, effects of cross-linguistic similarity are usually small, leading to the possibility that with a larger sample, a small effect of cross-linguistic similarity may become apparent. Indeed, previous studies and the BIA+ model would predict such an effect, though it is conceivable that the effect may be reduced by the lack of shared orthography. Future studies must revisit this issue with a larger number of participants and items, in order to confirm or disconfirm the findings presented here.

References

Allen, D. & Conklin, K. (2013). Cross-linguistic similarity and task demands in Japanese-English bilingual processing. *PLOS One, 8(8),* e72631. DOI:10.1371/journal.pone.0072631

Allen, D. & Conklin, K. (2014). Cross-Linguistic Similarity Norms for Japanese-English Translation Equivalents. *Behaviour Research methods, 46(2),* 540–563.

Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics.* London: Cambridge University Press.

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

Balling, L.W. (2013). Reading authentic texts: What counts as cognate? *Bilingualism: Language and Cognition, 16 (3),* 637-653.

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.M., & Maechler, M. (2010). *Lme4 Version 0.999375-37* (R Package). Last accessed on November 1, 2013 at: http://lme4.r-forge.r-project.org/

Bultena, S., Dijkstra, T., & van Hell, J. G. (2013). Cognate and word class ambiguity effects in noun and verb processing. *Language and Cognitive Processes, 28*(9), 1350–1377.

Bybee, J., & Hopper, P. (2001). *Frequency and the emergence of linguistic structure.* Amsterdam: Benjamins.

Cop, U., Keuleers, E., Drieghe, D., & Duyck, W. (2015). Frequency effects in monolingual and bilingual natural reading. *Psychonomic Bulletin & Review, 22 (5),* 1216-1234.

Cop, U., Dirix, N., Van Assche, E., Drieghe, D., & Duyck, W. (2016). Reading a book in one or two languages? An eye movement study of cognate facilitation in L1 and L2 reading. *Bilingualism, Language & Cognition,* (First View Article), DOI: http://dx.doi.org/10.1017/S1366728916000213.

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

Diependaele, K., Lemhöfer, K. & Brysbaert, M. (2013). The word frequency effect in first- and second-language word recognition: A lexical entrenchment account. *The Quarterly journal of experimental Psychology, 66 (5),* 843-863.

Dijkstra, A. (2007). The multilingual lexicon. In M.G. Gaskell, (Ed.), *Handbook of Psycholinguistics.* Oxford: Oxford University Press. (pp. 251-265)

Dijkstra, T., Grainger, J., & Van Heuven, W.J.B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41,* 496-518.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappeli, M., & Baayen, R.H. (2010). How cross-linguistic similarity affects cognate recognition. *Journal of Memory and Language, 62,* 284-301.

Dijkstra, T., & Van Heuven, W.J.B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5,* 175-197.

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective access. *Journal of Experimental Psychology: Language, Memory, and Cognition, 33,* 663-679.

Duyck, W., Vanderelst, D., Desmet, T., & Hartsuiker, R.J. (2008). The frequency effect in second-language visual word recognition. *Psychonomic Bulletin & Review 2008, 15 (4),* 850-85.

Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition, 18,* 91–126.

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24(2),* 143-188.

Ellis, N. C., R. Simpson-Vlach, and C. Maynard. (2008). Formulaic language in native and second-language speakers: psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly, 42(3),* 375–96.

Hino, Y., Lupker, S.J., & Pexman, P.M. (2002). Ambiguity and synonymy effects in lexical decision, naming and semantic 7 categorization tasks: Interactions between orthography, phonology and semantics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 686-713.

Hoshino, N., & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition, 106,* 501-511.

Kim, J., & Davis, C. (2003). Task effects in masked cross-script translation and priming. *Journal of Memory and Language, 49,* 484-499.

Libben, M., & Titone, D.A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 381-390.

McClelland, J.L., & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of Basic Findings. *Psychological Review, 88,* 375-407.

McDonald, S.A. & Shillcock, R.C. (2003). 'Eye-movements reveal the on-line computation of lexical probabilities during reading'. *Psychological Science, 14,* 648-52.

Odlin, T. (1989). Language Transfer: *Cross-Linguistic Influence in Language Learning.* Cambridge: Cambridge University Press.

Princeton University. (2010). "About WordNet". Available from: http://wordnet.princeton.edu

Rayner, K. (1998). Eye-movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372-422.

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

Ringbom, H. (1987). *The role of the first language in foreign language learning.* Clevedon: Multilingual Matters.

R Core Development Team. (2010) *R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing,* version 2.11.1. Available from: http://www.R-project.org.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language, 55*(2), 197–212.

Shibatani, M. (1990). *The Languages of Japan.* New York: 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(2),* 569–613.

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in sentence context. *Psychological Science, 20,* 923-927.

Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R.J. (2011). Does bilingualism change native-language reading? *Psychological Science, 20 (8),* 923-927.

Van Hell, J. G., & de Groot, A. M. B. (2008). Sentence context modulates visual word recognition and translation in bilinguals. *Acta Psychologica, 128*(3), 431–451.

Whitford, V., & Titone, D. (2012). Second-language experience modulates first- and second-language word frequency effects: Evidence from eye movement measures of natural paragraph reading. *Psychonomic Bulletin and Review, 19 (1),* 73–80.

Appendix 1

Cognate Targets

English	Japanese	Phonological Transcription	Phonological Similarity	Log word frequency (BNC)
bananas	バナナ	banana	3.4	6.1
bed	ベッド	beddo	4.2	7.3
bench	ベンチ	benchi	3.7	9.6
brush	ブラシ	burashi	3.1	7.4
classical	クラシック	kurashikku	3.7	7.9
coffee	コーヒー	koohii	3	8.7
curtains	カーテン	kaaten	3.4	7.1
diamond	ダイアモンド	daiamondo	3.5	6.5
flute	フルート	furuuto	3.8	5.7
gorillas	ゴリラ	gorira	3.3	4.3
guitars	ギター	gitaa	3.1	7.8
hammock	ハンモック	hanmokku	3.3	4.0
kangaroos	カンガルー	kangaruu	3.5	4.1
lions	ライオン	raion	3.7	6.5
pelicans	ペリカン	perikan	3.4	3.6
penguins	ペンギン	pengin	3.3	4.8
pipe	パイプ	paipu	3.7	7.7
plastic	プラスチック	purasuchikku	3.2	8.2
pool	プール	puuru	3.6	8.3
radio	ラジオ	rajio	2.5	8.7
shower	シャワー	shawaa	3.6	7.3
skirt	スカート	sukaato	3.2	7.2
taxi	タクシー	takushii	3.5	7.4
television	テレビ	terebi	2.4	9.1
trumpet	トランペット	toranpetto	3.7	5.8
violin	バイオリン	baiorin	3.3	6.2

Noncognate targets

English	Japanese	Phonological Transcription	Phonological Similarity	Log word frequency (BNC)
beautiful	美しい	utusukushii	1	8.98
bones	骨	hone	1	7.70
carrots	人参	ninjin	1	5.81
cloud	雲	kumo	1	7.57
dangerous	危険な	kikenna	1	8.60
destination	目的地	mokutekichi	1	7.00
ears	耳	mimi	1	7.89
effort	努力	douryoku	1	8.93
famous	有名な	yuumeina	1.2	8.73
favourite	好きな	sukina	1	8.45
finger	指	yubi	1	8.01
five	五	go	1	10.51
gentle	大人しい	otonashii	1	7.81
habits	習性	shusei	1	7.68
impressive	印象的	inshouteki	1.1	6.42
learn	学ぶ	manabu	1.1	8.98
lips	唇	kuchibiru	1	7.26
nose	鼻	hana	1.1	8.27
places	場所	basho	1	10.73
purchased	買う	kau	1	8.37
snakes	蛇	hebi	1.1	6.43
sun	太陽	taiyou	1.1	8.80
tree	木	ki	1	8.62
trousers	ズボン	zubon	1.1	5.67
zebras	シマウマ	shimauma	1	5.40
zoo	動物園	doubutsuen	1	5.89