[ACCEPTED MANUSCRIPT]

Parallel semantic processing in reading revisited: Effects of translation equivalents in bilingual readers

Joshua Snell, Mathieu Declerck & Jonathan Grainger

Laboratoire de Psychologie Cognitive, Aix-Marseille University & CNRS, Marseille, France

Short title: Parallel semantic processing in reading

Address for correspondence:
Joshua Snell
Laboratoire de Psychologie Cognitive,
Aix-Marseille University
3 Place Victor Hugo
13331 Marseille
France
E-mail: Joshua.snell@hotmail.com
Telephone: 04 13 55 09 99
Fax: 04 13 55 09 9

[Word count: 7675]

Abstract

Previous research has failed to establish semantic parafoveal-on-foveal effects during reading. As an explanation, we theorize that sentence reading engages a sentence-level representation that prevents semantic parafoveal-foveal integration. Putting this account to the test, we examined parafoveal-foveal influences both in- and outside a sentence reading setting. Optimizing chances of establishing parafoveal-on-foveal effects, we used translation-equivalent word pairs with French-English bilingual participants. Experiment 1 provided no evidence for semantic parafoveal-on-foveal integration during sentence reading, but some evidence that semantic information had been extracted in parallel from multiple words. Experiments 2 and 3 employed a flanker paradigm in which participants semantically categorized English foveal target words, while these were flanked by the French translation or an unrelated French word (stimulus on-time 170ms). Performance was drastically better with translation flankers, suggesting that readers can integrate semantic information across multiple words when the task does not require a strict separation of higher-order information.

Key words: Reading; Semantic parafoveal processing; Parafoveal-foveal integration; Translation equivalence; Bilingualism

While decades of reading research have yielded much insight into how foveal isolated words are processed and recognized, the mechanisms of parafoveal word processing are still far from understood. To gain understanding in the exact nature of parafoveal word processing, the field is required to answer two fundamental questions: (i) how deeply are parafoveal words processed, and (ii) how does the extraction of information from the parafovea influence foveal word processing and vice versa? It is unlikely that investigations into these matters will yield singular answers, as multiple lines of research have suggested that they are dependent on several factors, such as the language at hand (e.g. Deutsch, Frost, Pollatsek & Rayner, 2005; Bertram & Hyönä, 2007; Juhasz et al., 2009; Yan, Zhou, Shu & Kliegl, 2012; Schotter, Angele & Rayner, 2012 for a review) and inter-individual differences (e.g. Veldre & Andrews, 2014). Perhaps unsurprisingly then, the field has not yet come to a consensus, and years of research has divided scientists into, roughly, two schools of theorizing. One of these schools assumes that higher-order (e.g., lexical, semantic, syntactic) word processing occurs serially, i.e., for one word at a time, with attention moving from one word to the next when the former has been recognized (e.g. Angele, Tran & Rayner, 2013; Rayner, Schotter & Drieghe, 2014; Reichle, Pollatsek & Rayner, 2006). This means that if an upcoming word (*n+1*) can be lexically accessed at all, this would happen only after recognition of the foveal word (*n*), implying that the processing of word *n* should not be influenced by higher-order features of word *n+1*. The other school assumes that multiple words may be processed in parallel, as visuo-spatial attention would be distributed across multiple words as a gradient (Engbert, Nuthmann, Richter & Kliegl, 2005; Radach & Kennedy, 2004; Reilly & Radach, 2006). Consequently, it may be that the lexical properties of upcoming words are processed simultaneously with foveal words, and that higher-order features from upcoming words influence the foveal word recognition process.

For written languages that use an alphabetic script, there is considerable evidence that multiple words can be processed in parallel at the sub-lexical level. The principal finding is that foveal words are recognized faster when they are orthographically related to adjacent (upcoming) words – a so-called *parafoveal-on-foveal* effect (e.g. Angele et al., 2013; Dare & Shillcock, 2013; Grainger, Mathot & Vitu, 2014; Inhoff, Radach, Starr & Greenberg, 2000; Radach & Kennedy, 2004; Snell, Vitu &

Grainger, 2017a) – suggesting not only that upcoming words are processed to some extent prior to being fixated, but also that this happens *during* rather than *after* foveal word processing.

For higher-order parallel processing, matters seem to be more complicated. There is evidence that both semantic and syntactic information can be extracted from upcoming words prior to these words being fixated (see below); however, this is not direct evidence for parallel processing. Under the assumption of serial processing, for example, it is possible that higher-order processing of the upcoming word occurs during the interval in which the foveal word has been recognized but the eyes have not yet moved to the upcoming word (e.g. Schotter, Reichle & Rayner, 2014). It is important to note, however, that while evidence for higher-order parafoveal processing is compatible both with serial and parallel processing, an *absence* of such evidence would argue directly against parallel processing. Using the gaze-contingent boundary paradigm (Rayner, 1975) in German sentence reading, Hohenstein and Kliegl (2014) found that a target word (*n*) was recognized faster when there was a semantically related preview word at the target location when readers were fixating the pre-target (*n-1*). However, using the same paradigm, Rayner et al. (2014) could not establish such a semantic parafoveal preview benefit in English sentence reading. Similarly, Altarriba, Kambe, Pollatsek, and Rayner (2001) failed to find a preview benefit from translation equivalents during sentence reading by Spanish-English bilinguals. On the other hand, Schotter (2013) did find a semantic parafoveal preview effect in English sentence reading, but only when previews and targets were synonyms (e.g. *start – begin*), and not when they were associatively related (e.g. *ready – begin*), indicating that the likelihood of finding such an effect might depend on the strength of the relationship between words in some languages. More recent work by Veldre and Andrews (2016) further suggests that the plausibility of the preview word within the sentence context is a key factor, with semantic preview benefits being largely eradicated when the preview was not contextually plausible.

Finally, two recent studies have found evidence for syntactic processing of upcoming words. Snell, Meeter, and Grainger (2017b) found that target processing was facilitated when previews were syntactically congruent with the target (i.e., from the same syntactic category), as compared to when previews were from a different syntactic category. In a similar vein, Brothers and Traxler (2016) found that English readers were less likely to skip an upcoming word if it violated syntactic rules (e.g. noun followed by a noun).

While an absence of parafoveal preview effects would argue directly against parallel processing (and thus in favor of serial processing), higher-order *parafoveal-on-foveal* effects would argue directly against serial processing—the rationale being that processing of the upcoming word could only influence recognition of the fixated word if the latter was still being processed. Here, however, reports are again inconsistent. Snell et al. (2017a) showed that when word *n+1* was a high-frequency orthographic neighbor of word *n* (e.g. *rock rack*), processing of word *n* was facilitated, and not inhibited, as might be expected following lexical competition between orthographic neighbors (e.g. Davis & Lupker, 2006; Segui & Grainger, 1990). Furthermore, Angele et al. (2013) showed that processing of word *n* was not facilitated by a semantically related word *n+1* in English sentence reading. On the other hand, Inhoff et al. (2000) did find such a facilitatory effect in German sentence reading. Yet, as has been pointed out by Angele et al. (2013), it is unclear whether the effect reported by Inhoff et al. (2000) was purely semantic, or possibly orthographic in nature (e.g., processing of *'mother'* might be facilitated by '*father*' due to semantic relatedness, but also due to orthographic relatedness).

The study of Inhoff et al. (2000) taken aside, the general absence of higher-order parafoveal-on-foveal effects has been taken as evidence against parallel processing (e.g. Angele et al., 2013). However, here it should be noted that the premise that higher-order parafoveal-on-foveal effects evidence parallel processing, does not logically imply that an absence of such effects disproves parallel processing. As argued by Snell et al. (2017b), if multiple words are processed in parallel, it would be quite problematic if higher-order information is integrated across these words, given that each word has a distinct role in contributing to sentence comprehension. It would therefore seem more likely that if words are indeed processed in parallel, there would be a mechanism at play that allows readers to keep track of separate word identities, thus enabling independent extraction of semantic and syntactic information from words in the visual field. The mechanism that was theorized by Snell et al. (2017b) is a spatiotopic sentence-level representation in working memory, onto which activated lexical representations would be mapped (Figure 1). The location that an activated word is appointed in the sentence-level representation is determined by low-level visual cues (e.g. expectations about word length) and top-down grammatical constraints (e.g., 'I have recognized an article at position *n*, so I expect an adjective or noun at position *n+1*'). Similarly, top-down feedback from the sentence-level representation to individual words may attenuate or enhance their activation (e.g., the visual input '*This beer tastes good*' may activate both '*beer*' and '*been*', but the latter would be rejected due to grammatical constraints). Semantic and syntactic information associated with specific word identities at specific positions in the sentence would then provide the essential ingredients for computing sentence meaning. In sum, this account proposes that higher-order semantic and syntactic information can be processed in parallel across multiple words, but that the constraints associated with sentence reading ensure that such information is only integrated at the sentence level; hence the absence of higher-order parafoveal-on-foveal influences during sentence reading.

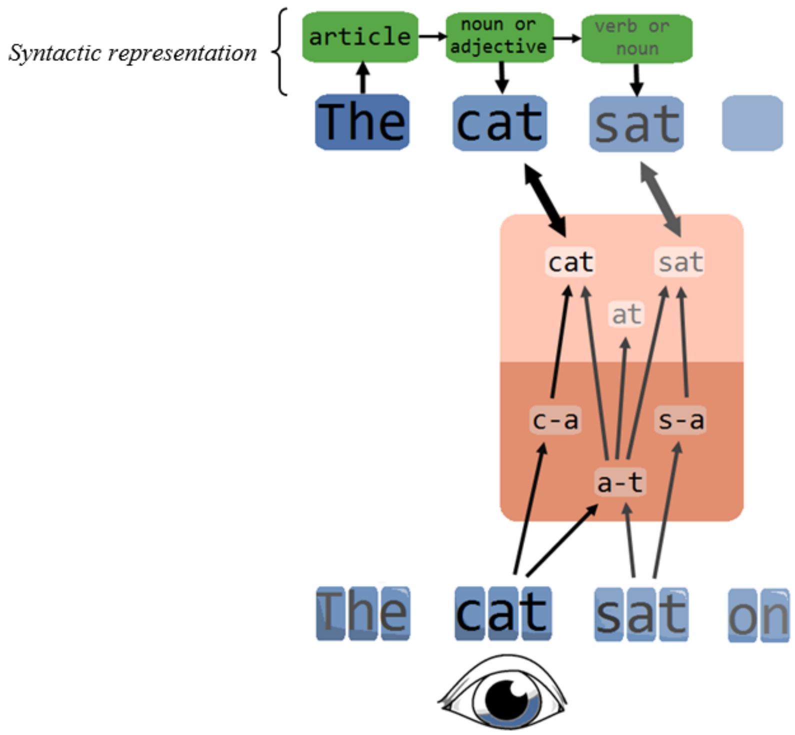

Figure 1. Our conceptualization of the reading system as proposed in Snell et al. (2017b). Sub-lexical orthographic information is gathered across multiple words, with stronger activation of letters in the fovea (here *'cat'*) than letters in the parafovea. Activated word representations are projected onto a plausible location in a spatiotopic representation, based on visual features such as word length and shape. From here, recognized words append to a sentence-level representation that follows syntactic rules: for instance, if word *n* is recognized as an article, word *n+1* is expected to be a noun or adjective (in English). Feedback from the syntactic level to the individual word positions constrains the recognition process for these words.

One key prediction of this model is that parafoveal-on-foveal effects as evidence for parallel processing should be observable in a paradigm that does not require sentence-level comprehension. In line with this prediction, we found that readers were faster and more accurate to classify Dutch foveal targets as noun or verb, when these were flanked by syntactically congruent flankers as compared with incongruent flankers (Snell et al., 2017b).

Further in line with the predictions of the model, in sentence reading, we found that syntactic information was extracted in parallel from– but not integrated across, the fixated and upcoming word: syntactically congruent word *n*+1's caused a tendency for *increased* rather than decreased fixation durations on word *n*, along with an increased fixation rate on word *n*+1 (Snell et al., 2017b). This is likely because the reading process was disturbed by readers' awareness of the grammatically incorrect continuation of the sentence (see also Snell et al., 2017a).

In the present study we provide a further test of higher-order parallel processing in both a sentence reading paradigm and a flanker paradigm, and this time for semantically rather than syntactically related parafoveal words, thus potentially unveiling differences in the natures of semantic parallel processing and syntactic parallel processing as investigated in Snell et al. (2017b). Indeed, the sentence-level feedback mechanism that was theorized here and in Snell et al. (2017b) mainly revolves around grammatical constraints, and it is possible that the behavioral patterns that were established in our previous study do not hold when using semantic relatedness.

We chose to test for effects of translation equivalents in bilingual participants, since translation equivalence arguably provides the strongest semantic relation between two words (e.g. Grainger & Frenck-Mestre, 1998; Perea, Duñabeitia & Carreiras, 2008; Duñabetia, Perea & Carreiras, 2010), thus maximizing the chances of obtaining semantic parafoveal-on-foveal effects. This provides a strong test of our model that predicts no such effect in sentence reading, accompanied by evidence for parallel processing of semantic information in the flanker paradigm.

Experiment 1

Methods

Participants

30 French-English bilingual students (11 female) from the Aix-Marseille Université (Marseille, France) gave written informed consent to their participation in this experiment. Participants earned €5 each for participating. All participants reported to be native to the French language, non-dyslexic and had normal or corrected-to-normal sight. The participants were naive with regard to the purpose of the experiment.

Prior to the actual experiment, the participant's proficiency in both French and English was tested with the *LexTALE* language proficiency test (Brysbaert, 2013; Lemhöfer & Broersma, 2012). From the initial 30 participants, 25 succeeded the English test with a score of at least 60%. We then recruited more participants to bring the total amount of proficient bilingual participants back to 30. The average score for French and English was 84.0% and 67.8% respectively.

	Organic	*Inorganic*
Target	wolf	hole
Translation	loup	trou
Control	loge	pion

Figure 2. Stimuli examples. As can be seen, if any of a target's constituent letters appear in its French translation word (underlined), these letters would also appear at the same position in the French control word.

Materials

We generated a set of 100 English target words with a length ranging between 3–6 letters. The French translation equivalent for each of these targets was a non-cognate word (e.g. *mist* – *brume*) with a length ranging between 3–7 letters. For every translation pair we retrieved a French control word from the French Lexicon Project database (Ferrand et al., 2010) that was similar to the translation word in terms of length and orthographic overlap with the target (see Figure 2). We also made sure that the average frequency of translations and controls was equal, at 4.72 and 4.82 Zipf, respectively.[1]

For every target we constructed an English sentence, fitting on a single line with length ranging between 23 and 47 characters (including spaces). Sentences contained 7.76 words on average (min. 5, max. 13 words). During stimulus presentation, we manipulated the post-target (n+1) prior to and during the fixation on the target (n), such that the post-target was either the French translation or the French control word. When participants moved their eyes from the target to the post-target, the latter would change into an English word that formed a logical continuation of the sentence (Figure 3). The average target/post-target boundary location was 15 characters from the start of the sentence (min. 8, max. 23 characters). For one out of every five sentences we also created a quiz question, shown directly after the sentence and to be answered with a two-button response. These served as catch trials to make sure that participants were reading for meaning.

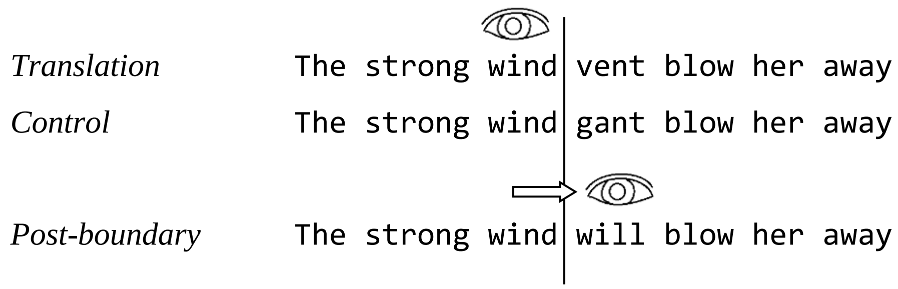

Figure 3. The upper two sentences show what a stimulus could look like in respectively the translation and control condition *before* the eyes crossed the invisible boundary (here marked by the vertical line). As soon as the eyes moved beyond the boundary, the post-target changed into a grammatically correct continuation of the sentence.

Design

Our experimental design consisted of two post-target word conditions (*translation / control*). A Latin-square design was used to ensure that all 100 sentences were presented in all conditions but only once per participant. The experiment thus consisted of 100 trials (50 with post-target translation word and 50 with post-target control word), and these were presented in randomized order.

---

1 For more on the Zipf frequency scale, see Van Heuven, Mandera, Keuleers and Brysbaert (2014).

Apparatus and software

The stimuli and experimental design were implemented with OpenSesame (Mathôt, Schreij & Theeuwes, 2012), with the PyGaze back-end (Dalmaijer, Mathôt & Van der Stigchel, 2014) to process eye movement data online. With an EyeLink 1000 (SR Research, Mississauga, ON, Canada), a video-based eye tracker sampling at 1000Hz with a spatial resolution of 0.01°, the reader's right eye position was recorded. Stimuli were presented on a 1024x768 px, 150 Hz computer monitor. Participants were seated at a distance of 90 cm from the display, so that each character space subtended 0.35 degrees of visual angle. Manual responses were collected with a keyboard. A chin-rest was used to stabilize the head position.

Procedure

Before commencing the experiment, the right eye was calibrated using a 3-point horizontal calibration grid with fixation points appearing in randomized order. In case of a sufficient match between the calibration grid and fixation grid, a validation was carried out to double-check the accuracy of the initial fixations. Participants then received instructions both verbally by the experimenter and visually on screen.

A drift correction was performed before the start of every trial. In case of a successful calibration, a forward slash (/) was presented as a fixation point, at a location that matched the start of the sentence when it appeared. As soon as the eyes had stabilized on the slash (within a 0.70 degrees range) for 700ms, the sentence stimulus appeared with the first letter aligned to the fixation location.

As participants read the sentence, the position of the eyes was tracked online. As soon as the eyes moved beyond an invisible boundary, the x-coordinate of which marked the exact middle between the target and the post-target word, the latter changed into a logical continuation of the sentence. When participants reached the end of the sentence, a green dot would briefly appear to the right of the sentence as a means to give recognition that the sentence was read. Shortly afterwards, the next trial would commence. However, if the sentence belonged to one of the 20 sentences for which we created a quiz question, participants would first see a display with that question and two possible answers in the left and right bottom corner of the screen. The participants had to choose one of these answers with respectively a left- or right-handed keyboard button response.

Participants were asked not to blink while reading the sentences but rather in-between the trials, because the temporary loss of corneal reflection could cause imprecise gaze estimations. The experiment lasted about 20 minutes and participants were allowed to take a break at their own convenience.

Results

From the total of 3000 trials across participants, 234 trials (7.80%) were discarded due to eye-blinking or a premature boundary change (due to landing too close to the boundary). We computed three fixation duration measures: the first fixation duration (FFD), gaze duration (GD) and total viewing time (TVT). Here, FFD refers to the first fixation on a target word, regardless of whether there is more than one fixation on this word. GD refers to the sum of all first-pass fixation durations on a target word, while TVT refers to the sum of all fixations on a target word, that is, including fixations after a regression. We also calculated three probability values: the probability that the target was skipped, the probability that the target was refixated during first-pass, and the probability that the target was refixated by means of an inter-word regression.

For the duration measures we used linear mixed-effect (LMM) models with items and subjects as crossed random effects (Baayen, 2008). The models were fitted with the lmer function from the lme4 package (Bates, Maechle, Bolker & Walker, 2015) in the R statistical computing environment. We report regression coefficients (*b*), standard errors (SE) and *t*-values for all factors. Fixed effects were deemed reliable if $|t| > 1.96$ (Baayen, 2008). Logistic LMM models (fitted with the glmer function) were used to analyze the skipping, refixation and regression probabilities. Here, the *z*-values can be interpreted in the same way as the *t*-values. In all analyses, values beyond 2.5 SD from the condition mean, (on average 2.06% of the trials), were marked as outliers and excluded.

Table 1. Mean fixation times and probabilities for Experiment 1.

	Condition means					
	FFD	*GD*	*TVT*	*Skip*	*Refix*	*Regress*
Translation	229.7 (87.7)	394.4 (217.1)	611.6 (425.9)	.08 (.08)	.47 (.16)	.40 (.14)
Control	233.2 (88.5)	393.4 (224.3)	614.8 (408.8)	.08 (.07)	.47 (.13)	.44 (.15)

Note: values in between parentheses indicate standard deviations. Abbreviations: FFD, first fixation duration; GD, gaze duration; TVT, total viewing time.

As it turned out, there were no significant effects in the fixation duration measures (Tables 1 and 2). The relatively low skipping rate, high refixation and regression rates and high GD and TVT values indicate that our participants' eye-movement behavior consisted of many fixations connected by short saccades, which, as has been argued by Rayner (1998), is typical of low-proficiency reading behavior.[2] The rate of regressions was significantly increased for the control condition as compared to the translation condition (Table 3). Although this may be an indication that comprehension of the target word was better in the translation condition, this was not reflected in the TVT. There were no significant effects in the other probability values.

Table 2. Analyses of fixation duration measures for Experiment 1 (ref: control).

	FFD			*GD*			*TVT*		
	b	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*
(Intercept)	233.06	6.44	**36.21**	390.57	18.59	**21.01**	624.74	41.46	**15.07**
Translation	-3.26	3.16	-1.03	5.26	7.51	0.70	-13.05	12.24	-1.07

Abbreviations: FFD, first fixation duration; GD, gaze duration; TVT, total viewing time; SE, standard error. Significant values are shown in bold.

[2] These behavioral patterns were not restricted to the target word: for instance, overall, 27% of the saccades were regressions in our study, whereas 10-15% would be typical of normal reading behavior (Rayner, 1998).

Table 3. Analysis of probability values for Experiment 1 (ref: control).

	Skip			*Refix*			*Regress*		
	b	*SE*	*z*	*b*	*SE*	*z*	*b*	*SE*	*z*
(Intercept)	-2.90	0.20	**-14.25**	-0.03	0.14	-0.23	0.04	0.23	0.16
Translation	0.00	0.14	-0.03	0.04	0.08	0.45	-0.26	0.09	**-2.86**

Note: significant values are shown in bold.

A Bayesian analysis of the null-hypothesis (i.e., that target words are not influenced by the semantic relatedness of post-target words in sentence reading) was carried out for the FFD, GD and TVT measures (Rouder, Speckman, Sun, Morey & Iverson, 2009). We found positive evidence favoring the null-hypothesis in all these measures, with BF_{01} = 3.17 for FFD, BF_{01} = 4.11 for GD, and BF_{01} = 3.05 for TVT (Kass & Raftery, 1995).

Lastly, we assessed whether the participant's L2 (English) proficiency had an impact on the (absence of an) effect of our manipulation. L2 proficiency was entered in a separate LMM as a 2-level factor, determined by whether the participant's LexTale score was above or below the group median. In none of the measures L2 proficiency modulated the effect of our manipulation (FFD: b = 5.80, SE = 6.43, t = 0.90; GD: b = 7.12, SE = 15.25, t = 0.47; TVT: b = 3.00, SE = 24.91, t = 0.12; skips: b = 0.04, SE = 0.30, z = 0.12; refixations: b = 0.11, SE = 0.17, z = 0.67; regressions: b = 0.14, SE = 0.19, z = 0.74). There was nonetheless a marginally significant main effect of L2 proficiency on the fixation duration (FFD: b = 21.62, SE = 12.02, t = 1.80). Interestingly however, the direction of this effect was such that greater proficiency led to longer fixation durations.

Discussion

Experiment 1 failed to find evidence for semantic parafoveal-on-foveal effects even in conditions where the semantic relation between the parafoveal and foveal word was maximal (i.e., translation equivalents). The strength of the semantic relation between translation equivalents is attested by research using the masked priming paradigm (Forster & Davis, 1984) showing robust effects of non-cognate translation primes in conditions where other types of semantic relation typically do not exhibit priming (e.g., Grainger & Frenck-Mestre, 1998; Perea et al., 2008; see Duñabetia et al., 2010, for a review). These translation priming effects are particularly robust when primes are in the first language (L1) and targets in the second language (L2), which corresponds to the parafoveal word in L1 and target in L2 in Experiment 1.

One could argue that the absence of a parafoveal-on-foveal influence of translation equivalents in Experiment 1 is due to the fact that participants were reading in a strictly monolingual context, and therefore that words from the other language would not be activated. Here, however, it is important to note that there is clear evidence that translation equivalents are activated in a strictly monolingual context (e.g., Thierry & Wu, 2007) in line with non-selective accounts of lexical access in bilinguals (e.g., Grainger & Dijkstra, 1992). Furthermore, the fact that there were fewer regressions in the translation condition than in the control condition suggests that semantic information was extracted from the parafoveal word, but that this information did not impact on foveal word processing, as predicted by our model.

Experiment 2 provides a test of the other side of the theoretical coin described in the Introduction. That is, that evidence for parallel semantic processing across multiple words should be observable in a task that does not require a strict separation of higher-order information. This task is the flanker task, with horizontally arranged flanker words placed left and right of a central target word, thus mimicking stimulus presentation in sentence reading. Crucially, and contrary to previous semantic flanker studies showing facilitatory effects of semantic relatedness (Shaffer & Laberge, 1979) and effects of translation equivalents in bilinguals (Guttentag, Haith, Goodman & Hauch, 1984)[3], target and flankers were presented very briefly (170 ms) to prevent the possibility that there would be enough time to process the flankers after the target was recognized, (i.e., there would only be time to process flankers *during* target processing). The short presentation time further prevented that any eye movements were made to the flanking stimuli. Using the same stimuli as in Experiment 1, we

---

[3] This prior research used vertically arranged flankers positioned above and below centrally located targets, and target and flankers remained on-screen until participants responded.

tested for effects of L1 translation equivalent flanker words on the semantic categorization of L2 target words in bilingual participants.

Experiment 2

Method

Participants

Twenty French-English bilingual students (12 female) from the Aix-Marseille University (Marseille, France) gave written informed consent to their participation in this experiment. Participants earned €4,- or its equivalent in course credit. All participants reported to be native to the French language, non-dyslexic, and had normal or corrected-to-normal vision. Further, all participants had *LexTALE* language proficiency test scores (Brysbaert, 2013; Lemhöfer & Broersma, 2012) of at least 60% (average scores for French and English were 86.6% and 65.0%, respectively).

Materials

The targets and post-targets from Experiment 1 were used as targets and flankers in Experiments 2 and 3. Of these targets, 50 corresponded to a natural object (e.g. *duck, neck, king*) and 50 corresponded to an artifactual object (e.g. *bridge, skirt, pen*).

Design

Our experimental design consisted of two flanker type conditions (*translation / control*). Every target was presented twice to each participant: once flanked by the translation, and once flanked by the control. The experiment thus consisted of 200 trials (100 with *natural* as correct response and 100 with *artifactual* as correct response), and these were presented in randomized order.

Apparatus and software

The stimuli and experimental design were implemented with OpenSesame (Mathôt, Schreij &

Theeuwes, 2012). Stimuli were presented on a 1024x768 px, 150 Hz computer monitor. Participants were seated at a distance of 90 cm from the display, so that each character space subtended 0.35 degrees of visual angle. Manual responses were collected with a keyboard.

Procedure

After taking the language proficiency test, participants received instructions both verbally by the experimenter and visually on screen. Every trial would start with two centrally positioned vertical fixation bars (see Figure 4). After 600 ms, the target word appeared in between these fixation bars, with the French translation or control word flanking its left and right side, (targets and flankers were separated by one character space on each side). After 170 ms, the target and flankers disappeared, and participants had 2000 ms to indicate whether they had recognized the target as being a natural or artifactual object. This was done with a left- or right-sided button press (‘w’ and ‘!’ respectively on an azerty layout keyboard), with the right button always corresponding to ‘natural’. A green or red dot was then briefly shown at the center of the screen, depending on whether the participant’s response was correct or incorrect respectively, shortly after which the next trial would commence. Before the start of the experiment, a set of twelve practice trials was run to allow participants to become acquainted with the procedure. A break was offered halfway through the experiment. The experiment lasted approximately 20 minutes in total.

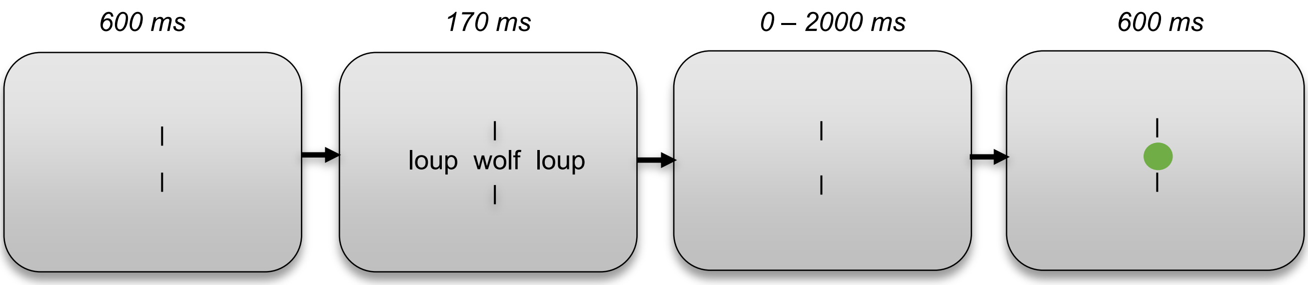

Figure 4. Overview of the trial procedure. The size of stimuli relative to the screen is exaggerated in this example.

Results

Trials where the response time (RT) was beyond 2.5 standard deviations from the condition mean (3.63% of all trials) were discarded. Only correctly answered trials (78.15% of all trials) were included in the RT analyses. For our RT analyses we again used LMMs with items and subjects as crossed random effects, fitted with the lmer function from the lme4 package (Bates et al., 2015) in the R statistical computing environment. We report regression coefficients (*b*), standard errors (SE) and *t*-values. Logistic LMMs (fitted with the glmer function) were used to analyze the error rates.

Table 4 shows the mean RTs and error rates for the translation and control condition. RTs were significantly lower in the translation condition than in the control condition, with b = 40.25, SE = 7.17, t = 5.62. This effect of flanker type did not differ significantly between trials where the target was a natural object and trials where the target was an artifactual object: b = 18.36, SE = 14.34, t = 1.28. The error rate was significantly lower in the translation condition as well, with b = 0.40, SE = 0.09, z = 4.41.

Table 4. Mean RT's and error rates for the translation and control condition of Experiment 2.

	RT	*Error*
Translation	625.03 (232.04)	.166 (.060)
Control	662.91 (246.38)	.216 (.071)

Note: values in between parentheses indicate standard deviations.

Discussion

In Experiment 2, we set out to investigate whether readers can semantically process multiple words in parallel. To maximize our chances of finding evidence for such parallel processing of semantic information, we used non-cognate translation equivalents with bilingual participants, as in Experiment

1. The results of Experiment 2 suggest that readers indeed process semantic information across foveal and parafoveal stimuli, as target processing was strongly facilitated by translation flankers, compared to control flankers. This effect could not have been caused by sub-lexical factors such as orthographic overlap, as the orthographic overlap (letter identity and position) with the target was identical for translations and controls. Moreover, the stimulus presentation time was considerably short (170 ms), suggesting that there would not have been enough time to process the flankers after the target was recognized, as would otherwise be possible under the assumption of serial processing. The size of the effect (b = 40.25) is considerably larger than those found in previous implementations of this paradigm, such as an earlier study in which we used syntactically related stimuli (b = 17.45; Snell et al., 2016b) and studies that examined sub-lexical parafoveal-on-foveal effects (Dare & Shillcock, 2013; Grainger et al., 2014; Snell et al., 2016a; all b-values <20.00), indicating that stronger higher-order relations between foveal and parafoveal words may indeed lead to greater parafoveal-on-foveal influences.

The primary aim of Experiment 3 was to replicate the results of Experiment 2, while including a third *no-flanker* condition. According to our model, attentional resources are distributed across multiple words. As a result, orthographically unrelated flanker stimuli, such as the translation and unrelated flankers used in Experiment 2, should interfere with the orthographic processing of the target. This should lead to increased difficulty in target processing in the presence of flanking words compared with a no-flanker condition.

Experiment 3

Method

The methodology for Experiment 3 was nearly identical to that of Experiment 2, the only difference being the inclusion of a third, *no-flanker* condition. The 100 targets from Experiment 2 were also used in Experiment 3. This time, each target was presented three times to every participant, corresponding to the three flanker conditions, thus making the total amount of experimental trials 300. We again recruited twenty students (15 female), none of whom had participated in the previous experiments. This group of participants had average *LexTALE* language proficiency scores of 92.0% and 63.8% for French and English, respectively. The experiment lasted approximately 25 minutes.

Results

We applied criteria identical to those used in Experiment 2 for the exclusion of trials in Experiment 3. We ended up with data from 79.87% of all trials for the analysis of RTs. We again employed LMMs for the analyses of RTs and error rates. Condition means for Experiment 3 are shown in Table 5. Importantly, we replicated our results from Experiment 2, as RTs in the translation condition were again significantly decreased as compared to the control condition: b = 35.59, SE = 6.87, t = 5.18. We also hypothesized that orthographically unrelated flanker words should interfere with the orthographic processing of target words hence slowing target word processing compared with the no-flanker condition. This was indeed the case. The no-flanker condition yielded even lower RTs than the translation condition: b = 28.83, SE = 6.84, t = 4.22. As in Experiment 2, the flanker effect did not differ between trials where the target was a natural object and trials where the target was an artifactual object: b = 9.01, SE = 6.89, t = 1.31. Although the error rate was again numerically lower in the translation condition than in the control condition, the difference did not reach significance this time around: b = 0.14, SE = 0.09, z = 1.46. The error rates did not differ between the translation and no-flanker condition either: b = 0.10, SE = 0.09, z = 1.06.

Table 5. Mean RT's and error rates for the no-flanker, translation and control condition of Experiment 3.

	RT	*Error*
No-flanker	608.44 (219.50)	.178 (.081)
Translation	638.95 (224.08)	.168 (.073)
Control	672.47 (232.76)	.179 (.092)

Note: values in between parentheses indicate standard deviations.

Discussion

We replicated our results from Experiment 2 in Experiment 3. Together, these experiments provide clear evidence that readers cannot effectively focus their attention on single words, causing higher-order processing to take place in parallel for foveal and parafoveal words. The fact that the no-flanker condition yielded even lower RTs than the translation condition is in line with our model, according to which orthographically unrelated flankers will interfere with target word processing. Thus, in the absence of grammatical constraints, the inevitable processing of multiple words leads to cross-leakage of information both at the sub-lexical level, where orthographically unrelated flanking stimuli (i.e., both the non-cognate translation equivalents and the unrelated control words) interfere with orthographic processing of the target word, as well as beyond the lexical level, where semantic information extracted from the translation equivalents facilitates semantic categorization.

General discussion

In the present study we examined semantic parafoveal-on-foveal effects by testing bilingual participants with parafoveal words that were translation equivalents of the foveal target word. Our prior research investigating syntactic parafoveal-on-foveal effects (Snell et al., 2017b) pointed to parallel independent extraction of syntactic information from multiple words during sentence reading. The independent nature of such parallel processing, induced by top-down grammatical constraints, was taken as the explanation for why no syntactic parafoveal-on-foveal effects were observed. However, when integration of syntactic information was beneficial for the task at hand, we found evidence for parallel processing of syntactic information. This task involved syntactic classification of a central target word flanked by unrelated flanker words that could be from the same or a different syntactic category. Syntactic information extracted from multiple words in parallel could then be pooled into a single response channel, hence the facilitation from syntactically congruent flankers. The present study built directly on this prior work, but now testing for semantic parafoveal-on-foveal effects, and using what is arguably the strongest semantic relation between two words – translation equivalence. We predicted an absence of semantic parafoveal-on-foveal effects in sentence reading, accompanied by evidence for parallel processing of semantic information in the flanker paradigm.

Experiment 1 employed a parafoveal-on-foveal manipulation with the boundary technique during sentence reading. The presence of a translation equivalent at position *n*+1 was found to have no significant influence on the processing of word *n*, as revealed by fixation duration measures. Some evidence for parallel semantic processing of words *n*+1 and *n* was found, however, in the form of a decreased rate of regressive saccades to word *n* when *n*+1 was a translation of *n*. Yet, the decreased regression rate was not reflected in the total viewing time (TVT) on the target, suggesting that parallel semantic processing may influence higher levels of reading comprehension, rather than processing of individual words. Furthermore, given the recent results of Veldre and Andrews (2016), the absence of a semantic parafoveal-on-foveal effect in Experiment 1 could be due to the fact that the translation equivalent did not fit plausibly into the sentence context – a scenario that fits well with how the spatiotopic sentence-level representation (Figure 1; Snell et al., 2017b) would operate.

Experiments 2 and 3 used a flanker paradigm with horizontally arranged flanker words located to the left and to the right of a central target word. Flanker words could be the translation equivalent in L1 of the target word in L2, or completely unrelated L1 words. Semantic categorization of target words was facilitated by the translation flankers relative to the unrelated flankers. The stimulus presentation time was considerably short (170 ms), such that there would not have been enough time to process the flankers after recognition of the target, as would otherwise be possible if words were processed serially rather than in parallel. Furthermore, in Experiment 3 we found that target identification was faster when there were no flanker stimuli compared with the translation flanker condition. This is captured in our model by the interference generated from orthographically unrelated flanking stimuli, following an attentional distribution spanning multiple words. Semantically related flankers help reduce this flanker interference by providing congruent semantic input into the mechanism that decides whether the target word is a living thing or not.

One might wonder why Altarriba et al. (2001) failed to find an influence of parafoveal translation equivalents in a paradigm (their Experiment 1) that shares certain similarities with the flanker paradigm used in Experiments 2 and 3 of the present study. In the Altarriba et al. study, bilingual participants had to fixate a central fixation cross while a parafoveal word was presented at 2° of eccentricity (fixation cross to beginning of word) to the right of fixation. Participants made an eye movement to the parafoveal stimulus, and during that eye movement the preview word was replaced with the target word that participants had to read aloud as rapidly as possible. One major difference with respect to our flanking paradigm is that the parafoveal word and foveal word were presented sequentially at the same location in Altarriba et al.'s study, as opposed to the parallel presentation at different locations in our study. Thus, any benefit of a semantically related preview in the Altarriba et al. study might have been cancelled by the interference caused by having orthographically different stimuli appear at the same location (see also e.g. Kliegl, Hohenstein & McDonald, 2013; Marx, Hawelka, Schuster & Hutzler, 2017, regarding the interplay of preview benefit versus preview cost).

Overall, the results of the present study are in line with the predictions of a new model of parallel word processing and reading (Snell et al., 2017b). According to this model, orthographic information spanning several words is integrated into a single processing channel (Grainger et al., 2014; 2016), hence explaining orthographic parafoveal-on-foveal effects. Orthographic word identities continue to be processed in parallel, nevertheless, with each word identity being associated with a particular position in the sentence that is being read. This parallel, independent, location-specific processing of word identities enables parallel independent activation of semantic and syntactic information from multiple words, which then feed information into higher-level sentence comprehension processes (Figure 1). The independent nature of word-level processing means that neither semantic nor syntactic parafoveal-on-foveal effects should be observed. On the other hand, in paradigms where this information can be pooled in order to generate a response, one can demonstrate parallel processing of semantic and syntactic information across multiple words.

To finish on a methodological note, a common criticism of reading research using static paradigms such as the flanker paradigm, is that these paradigms do not reflect normal reading. In response to this, we would point out that such paradigms provide theoretical leverage that cannot be achieved in sentence reading paradigms, and that what is crucial here is the possibility to create fundamental connections between processing involved in the simplified paradigms and processing involved in the more complex, and naturalistic, sentence-reading paradigms. Our model of parallel word processing allows us to establish such connections and to use the data obtained from multiple paradigms to inform the general mechanisms involved in everyday reading. Crucially, the flanker paradigm shows that readers are not able to effectively focus attention on single words. While it is evident that the flanker paradigm is different from sentence reading, serial processing accounts are challenged by the question of how readers would then be better at focusing attention on single words during sentence reading, given that (i) the visual input during sentence reading is more complex than in the flanker paradigm, and more dynamic due to eye-movements, (ii) parafoveal information is of interest during sentence reading, and (iii) parafoveal information is available longer during sentence reading.

Finally, the theoretical framework that is discussed in this paper may be tested in ways other than those employed here. For instance, Snell and Grainger (2017) have recently employed a paradigm that seems to sit neatly in between the flanker paradigm and natural sentence reading. It was demonstrated that the identification of a target word in a briefly presented sequence of four words is facilitated when the words form a grammatically correct sentence, compared to when the same words are presented in a shuffled agrammatical order. Future research may further endeavor to create a 'sentence reading' setting in which readers do not engage the sentence-level representation: for instance, higher-order parafoveal-on-foveal influences may be found during the reading of random, agrammatical word sequences.

In sum, results from the present experiments suggest that readers can extract semantic information from multiple words at once. In sentence-reading, this information is appended to a sentence-level representation rather than integrated as a whole (Figure 1), explaining why parafoveal-on-foveal effects in sentence reading have been elusive. As seen in Experiments 2 and 3, however, higher-order information can be integrated across multiple words when readers do not engage this sentence-level representation.

Acknowledgements

The authors would like to thank Margot Roisin and Svenja Krabbe for their help in conducting this research. This research was funded by the French National Research Council (ANR-11-LABX-0036, ANR-15-CE33-0002-01).

References

Altarriba, J., Kambe, G., Pollatsek, A. & Rayner, K. (2001). Semantic codes are not used in integrating information across eye fixations in reading: Evidence from fluent Spanish-English bilinguals. *Perception & Psychophysics, 63*, 875–890. doi: 10.3758/BF03194444

Angele, B., Tran, R. & Rayner, K.(2013). Parafoveal–foveal overlap can facilitate ongoing word identification during reading: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 39*, 526–538. doi: 10.1037/a0029492

Baayen, R. (2008). *Analyzing Linguistic Data: A pratical introduction to statistics*. Cambridge: Cambridge University Press.

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2015), Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67, 1–48. doi: 10.18637/jss.v067.i01

Bertram, R. & Hyönä, J. (2007). The interplay between parafoveal preview and morphological processing in reading. In van Gompel, R., Fischer, M., Murray, W. & Hill, R. (Eds.). *Eye movements: A window on mind and brain*. Oxford, UK: Elsevier. doi: 10.1075/ml.6.1.04ber

Brothers, T., & Traxler, M. (2016). Anticipating syntax during reading: Evidence from the boundary change paradigm. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *42*, 1894–1906. doi: 10.1037/xlm0000257

Brysbaert, M. (2013). LEXTALE_FR: A fast, free, and efficient test to measure language proficiency in French. *Psychologica Belgica, 53*, 23–37. doi: 10.5334/pb-53-1-23

Dalmaijer, E., Mathôt, S. & Van der Stigchel, S. (2014). PyGaze: An open-source, crossplatform toolbox for minimal-effort programming of eyetracking experiments. *Behavior Research Methods, 46*, 913–921. doi: 10.3758/s13428-013-0422-2

Dare, N. & Shillcock, R. (2013). Serial and parallel processing in reading: Investigating the effects of parafoveal orthographic information on nonisolated word recognition. *Quarterly Journal of Experimental Psychology, 66*, 417–428. doi: 10.1080/17470218.2012.703212

Davis, C. & Lupker, S. (2006). Masked inhibitory priming in English: Evidence for lexical inhibition. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 668–687. doi: 10.1037/0096-1523.32.3.668

Deutsch, A., Frost, R., Pollatsek, A. & Rayner, K. (2005). Morphological parafoveal preview benefit effects in reading: Evidence from Hebrew. *Language and Cognitive Processes*, *20*, 341–371. doi: 10.1080/01690960444000115

Duñabetia, J.A., Perea, M. & Carreiras, M. (2010). Masked translation priming effects with highly proficient simultaneous bilinguals. *Experimental Psychology*, *57*, 98-107. doi: 10.1027/1618-3169/a000013

Engbert, R., Nuthmann, A., Richter, E. & Kliegl, R. (2005). SWIFT: A dynamic model of saccade generation during reading. *Psychological Review, 112,* 777–813. doi: 10.1037/0033-295X.112.4.777

Ferrand, L., et al. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods, 42,* 488–496. doi: 10.3758/BRM.42.2.488

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

Grainger, J. & Dijkstra, T. (1992). On the representation and use of language information in bilinguals. In R.J. Harris (Ed.) *Cognitive processing in bilinguals*. Amsterdam: North Holland.

Grainger, J., Dufau, S. & Ziegler, J.C. (2016). A vision of reading. *Trends in Cognitive Sciences*, *20*, 171-179. doi: 10.1016/j.tics.2015.12.008

Grainger, J. & Frenck-Mestre, C. (1998). Masked translation priming in bilinguals. *Language and Cognitive Processes, 13,* 601–623. doi: 10.1080/016909698386393

Grainger, J., Mathôt, S. & Vitu, F. (2014). Test of a model of multi-word reading: Effects of parafoveal flanking letters on foveal word recognition. *Acta Psychologica*, *146*, 35–40. doi: 10.1016/j.actpsy.2013.11.014

Guttentag, R., Haith, M., Goodman, G. & Hauch, J. (1984). Semantic processing of unattended words by bilinguals: A test of the input switch mechanism. *Journal of Verbal Learning and Verbal Behavior, 23,* 178–188. doi: 10.1016/S0022-5371(84)90126-9

Hohenstein, S. & Kliegl, R. (2014). Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 166–190. doi: 10.1037/a0033670

Hyönä, J. & Häikiö, T. (2005). Is emotional content obtained from parafoveal words during reading? An eye movement analysis. *Scandinavian Journal of Psychology, 46*, 475–483. doi: 10.1111/j.1467-9450.2005.00479

Inhoff, A., Radach, R., Starr, M. & Greenberg, S. (2000). Allocation of visuospatial attention and saccade programming during reading. In Kennedy, A., Radach, R., Heller, D. & Pynte, J. (Eds.). *Reading as a perceptual process*. Oxford, UK: Elsevier.

Juhasz, B., Pollatsek, A., Hyönä, J., Drieghe, D. & Rayner, K. (2009). Parafoveal processing within and between words. *Quarterly Journal of Experimental Psychology, 62,* 1356–1376. doi: 10.1080/17470210802400010

Kass, R. & Raftery, A. (1995). Bayes Factors. *Journal of the American Statistical Association, 90,* 773–795. doi: 10.1080/01621459.1995.10476572

Kliegl, R., Hohenstein, S. & McDonald, S. (2013). How preview space/time translates into preview cost/benefit for fixation durations during reading. *Quarterly Journal of Experimental Psychology, 66,* 581-600. doi: 10.1080/17470218.2012.658073

Lemhöfer, K. & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods, 44,* 325–343. doi: 10.3758/s13428-011-0146-0

Marx, C., Hawelka, S., Schuster, S. & Hutzler, F. (2017). Foveal processing difficulty does not affect parafoveal preprocessing in young readers. *Scientific Reports, 7,* e41602. doi: 10.1038/srep41602

Mathôt, S., Schreij, D. & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*, 314–324. doi: 10.3758/s13428-011-0168-7

Perea, M., Duñabeitia, J. & Carreiras, M. (2008). Masked associative/semantic priming effects across languages with highly proficient bilinguals. *Journal of Memory and Language, 58,* 916–930. doi: 10.1016/j.jml.2008.01.003

Radach, R. & Kennedy, A. (2004). Theoretical perspectives on eye movements in reading: Past controversies, current issues, and an agenda for future research. *European Journal of Cognitive Psychology, 16,* 3–26. doi: 10.1080/09541440340000295

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65–81. doi: 10.1016/0010-0285(75)90005-5

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–422. doi: 10.1.1.294.4262

Rayner, K., Schotter, E. & Drieghe, D. (2014). Lack of semantic parafoveal preview benefit revisited. *Psychonomic Bulletin & Review, 21,* 1067–1072. doi: 10.3758/s13423-014-0582-9

Reichle, E., Pollatsek, A. & Rayner, K.(2006). E-Z Reader: A cognitive-control, serial-attention model of eye movement behavior during reading. *Cognitive Systems Research*, *7,* 4–22. doi: 10.1016/j.cogsys.2005.07.002

Reilly, R. & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research, 7,* 34–55. doi: 10.1016/j.cogsys.2005.07.006

Rouder, J., Speckman, P., Sun, D. & Morey, R. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16,* 225–237. doi: 10.3758/PBR.16.2.225

Schadder, W. & Laberge, D. (1979). Automatic semantic processing of unattended words. *Journal of Verbal Learning and Verbal Behavior, 18,* 413–426. doi: 10.1016/S0022-5371(79)90228-7

Schaffer, W., & LaBerge, D. (1979). Automatic semantic processing of unattended words. *Journal of Verbal Learning & Verbal Behavior*, *18,* 413–426. doi: 10.1016/S0022-5371(79)90228-7

Schotter, E. (2013). Synonyms provide semantic preview benefit in English. *Journal of Memory and Language, 69,* 619–633. doi: 10.1037/a0036763

Schotter, E., Angele, B. & Rayner, K. (2012). Parafoveal processing in reading. *Attention Perception & Psychophysics, 74,* 5–35. doi: 10.3758/s13414-011-0219-2

Schotter, E., Reichle, E. & Rayner, K. (2014). Rethinking parafoveal processing in reading: Serial attention models can account for semantic preview benefit and n+2 preview effects. *Visual Cognition, 22*, 309–333. doi: 10.1080/13506285.2013.873508

Schwanenflugel, P. & Rey, M. (1986). Interlingual semantic facilitation: Evidence for a common representational system in the bilingual lexicon. *Journal of Memory and Language, 25*, 605–618. doi: 10.1016/0749-596X(86)90014-8

Segui, J., & Grainger, J. (1990). Priming word recognition with orthographic neighbors: Effects of relative prime-target frequency. *Journal of Experimental Psychology: Human Perception and Performance*, *16,* 65–76. doi: 10.1037/0096-1523.16.1.65

Snell, J. & Grainger, J. (2017). The sentence superiority effect revisited. *Cognition, 168,* 217-221. doi: 10.1016/j.cognition.2017.07.003

Snell, J., Vitu, F. & Grainger, J. (2017a). Spatial integration of parafoveal orthographic information: Beyond the sub-lexical level? *Quarterly Journal of Experimental Psychology, 70,* 1984–1994. doi: 10.1080/17470218.2016.1217247

Snell, J., Meeter, M. & Grainger, J. (2017b). Evidence for simultaneous syntactic processing of multiple words during reading. *PLoS ONE, 12,* e0173720. doi:10.1371/journal.pone.0173720

Thierry, G. & Wu, Y. J. 2007. Brain potentials reveal unconscious translation during foreign language comprehension. *Proceeding of National Academy of Sciences*, *104*, 12530–12535. doi: 10.1073/pnas.0609927104

Van Heuven, W., Mandera, P., Keuleers, E. & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology, 67*, 1176–1190. doi: 10.1080/17470218.2013.850521

Veldre, A. & Andrews, S. (2014). Parafoveal preview benefit is modulated by the precision of skilled readers' lexical representations. *Journal of Experimental Psychology: Human Perception & Performance, 41,* 219–232. doi: 10.1037/xhp0000017

Yan, M., Zhou, W., Shu, H. & Kliegl, R. (2012). Lexical and sublexical semantic preview benefits in Chinese reading. *Journal of Experimental Psychology: Learning, Memory and Congnition, 38,* 1069–1075. doi: 10.1037/a0026935