

Parafoveal Processing Within and Between Words

Barbara J. Juhasz

Wesleyan University

Alexander Pollatsek

University of Massachusetts, Amherst

Jukka Hyönä

Turku University

Denis Drieghe

Ghent University

and

Keith Rayner

University of California, San Diego

Corresponding Author:

Barbara J. Juhasz

Department of Psychology

207 High Street

Wesleyan University

Middletown, CT 06424

Phone: 860-685-4978

Email: bjjuhasz@wesleyan.edu

Abstract

Parafoveal preview was examined within and between words in two eye movement experiments. In Experiment 1, unspaced and spaced English compound words were used (e.g., *basketball*, *tennis ball*). Prior to fixating the second lexeme, either a correct or a partial parafoveal preview (e.g., *ball* or *badk*) was provided using the boundary paradigm (Rayner, 1975). There was a larger effect of parafoveal preview on unspaced compound words than on spaced compound words. However, the parafoveal preview effect on spaced compound words was larger than would be predicted on the basis of prior research. Experiment 2 examined whether this large effect was due to spaced compounds forming a larger linguistic unit by pairing spaced compounds with non-lexicalized adjective-noun pairs. There were no significant interactions between item type and parafoveal preview suggesting that it is the syntactic predictability of the noun which is driving the large preview effect.

According to serial attention shift (SAS) models of eye movements in reading, lexical processing is the engine that drives the eyes forward in text (e.g. E-Z Reader, Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Pollatsek, & Rayner, 2003) and words are identified serially, one at a time. This view differs from parallel lexical processing models such as SWIFT (Engbert, Nuthmann, Richter, & Kliegl, 2005) which posit that multiple words can be processed in parallel as long as they are within a gradient of attention. Thus, to accurately describe the reading process requires a good understanding of how words are recognized. There is a large literature detailing how words are identified both in isolation (e.g. Rayner & Pollatsek, 1989) and during normal silent reading (see Rayner, 1998 for a review). Until relatively recently, however, most research has focused on relatively short monomorphemic words, and thus most models of word recognition were limited in their scope.

While there is increasing interest in how morphologically complex words are recognized, the majority of data used to inform models of word recognition come from experiments where words are presented in isolation. These word-in-isolation tasks impose task demands not encountered in normal silent reading. They also ignore the role of sentence context in word recognition (as often no context is provided), and do not take into consideration the effect of parafoveal processing on a word before it is fixated, as most words are presented in the fovea. Visual acuity is best in the fovea and letters in this region can be easily identified, but the foveal region only extends from the point of fixation to 1 degree of visual angle in either direction horizontally. (In reading, 1 degree of visual angle typically corresponds to 3-4 character spaces in an alphabetic language.) The parafoveal region extends outside the foveal region up to about 5 degrees in both horizontal directions, and in it, visual acuity drops off markedly with the distance from the fovea. The primary purpose of eye movements in reading is to bring to-be-read

text into the fovea where it can be easily identified. Experiments on reading have demonstrated that the perceptual span extends roughly 15 characters to the right in English but only 3-4 characters to the left (McConkie & Rayner, 1975; 1976; Rayner & Bertera, 1979; Rayner, Well, & Pollatsek, 1980). Thus, information is typically gained during a fixation not just from the currently fixated word, but also from a word or two to the right of fixation.

The types of information that are obtained from a parafoveal word prior to it being fixated have been identified using an eye movement paradigm called the *boundary paradigm* (Rayner, 1975). In this paradigm an invisible boundary is specified immediately before a target word in a sentence, and prior to fixation, a parafoveal preview of the target word is provided. Once the readers move their eyes across the boundary, the preview is replaced by the actual target word. Since this change occurs during an eye movement, when vision is functionally suppressed, readers are not aware of the change. Yet, reading times on the target word are influenced by the nature of the parafoveal preview. Research has revealed that readers gain information about the first few letters of a parafoveal word prior to fixation (Rayner, 1975; Rayner, McConkie, & Zola, 1980). They also gain information about the sound of the parafoveal word (Ashby, Treiman, Kessler, & Rayner, 2006; Chace, Rayner, & Well, 2005; Henderson, Dixon, Peterson, Twilley, & Ferreira, 1995; Pollatsek, Lesch, Morris, & Rayner, 1992) and the length of the parafoveal word (Inhoff, Radach, Eiter, & Juhasz, 2003; Juhasz, White, Liversedge, & Rayner, 2008; White, Rayner, & Liversedge, 2005b). However, research has failed to demonstrate that readers obtain any partial semantic information from parafoveal words (Altarriba, Kambe, Pollatsek, & Rayner, 2001; Hyönä & Häikiö, 2005; Rayner, Balota, & Pollatsek, 1985), and research in alphabetic languages has failed to demonstrate that readers gain

information about the morphological status of the parafoveal word (Bertram & Hyönä, 2007; Kambe, 2004; Lima, 1987).

The present study is concerned with whether parafoveal information from the end of a word being fixated is utilized in the same manner as parafoveal information from the next word, and what the implications are of this question for morphological processing. In order to examine this issue, we must first examine what constitutes a ‘word’ in the English language. To native English speakers, this may seem like an easy question because words in written English are typically separated by blank interword spaces. As a result, most people would define a word orthographically: as a sequential series of letters with a space before and after, and which has a meaning attached to it. Thus, *cat*, *dog*, *house*, and *pet* would all be classified as words, as well as the morphologically complex compounds *cathouse* and *doghouse*. However, to most native speakers, the phrase *pet dog* in the sentence “My pet dog barked” would not constitute a word because the blank space would define *pet dog* as two separate words. The story of what makes a word a word is more complex in other languages, such as Chinese, where interword spaces are not used (see, for example Bai, Yan, Liversedge, Zang, & Rayner, 2008). Even in English, however, there is a linguistic gray area regarding words. A class of items exists which have been referred to as “spaced compound words” (Juhasz, Inhoff, & Rayner, 2005). These are two-word expressions such as *front door* and *rush hour* which refer to a single concept, are familiar to native speakers, and appear in dictionaries. In fact, there are no specific rules for compounding in the English language (consider *basketball* and *tennis ball*). In contrast, some other alphabetic languages, such as Finnish, Dutch, and German, have clear rules for compounding, and in these languages, spaced compounds simply do not exist (even for novel compounds).

There is now a considerable literature on the processing of unspaced compound words during silent reading – both in very productive compounding languages such as Finnish (Bertram & Hyönä, 2003; Hyönä, Bertram & Pollatsek, 2004; Hyönä & Pollatsek, 1998; Pollatsek & Hyönä, 2005; Pollatsek, Hyönä, & Bertram, 2000) and in less productive compounding languages such as English (Andrews, Miller, & Rayner, 2004; Juhasz, Starr, Inhoff, & Placke, 2003; Juhasz, 2007). The research in both types of languages has provided evidence that familiar unspaced compounds are decomposed into their two words (henceforth lexemes) during reading. The key manipulation in these experiments was varying the frequency of either the beginning or ending lexeme in the compound words while holding the overall compound word frequency constant. In reading, high frequency words are processed faster than low frequency words and this is reflected in fixation durations on the words (Inhoff & Rayner, 1986; Rayner & Duffy, 1986; Schilling, Rayner & Chumbley, 1998). Compound words with a high frequency beginning or ending lexeme are fixated for a shorter duration than compound words with a low frequency lexeme (Andrews et al., 2004; Bertram & Hyönä, 2003; Hyönä et al., 2004; Hyönä & Pollatsek, 1998; Juhasz et al., 2003; Juhasz, 2007; Pollatsek & Hyönä, 2005; Pollatsek et al., 2000) indicating that accessing the lexemes of the compound is a functional aspect of whole compound recognition. The beginning lexeme is accessed before the ending lexeme (though the beginning lexeme frequency effect is more elusive in English). This appears to be especially true when the compound words are long in Finnish (Bertram & Hyönä, 2003). One may thus be tempted to conclude that unspaced compound words are processed in the same way as two separate words in text and thus should not be thought of as a single word. However, further research with Finnish has revealed an effect of the frequency of the whole compound word that has the same time-

course as the second lexeme frequency effect in that language, suggesting that a whole word representation also exists for unspaced compounds (Pollatsek et al., 2000).

Also in support of the “wordness” of compounds, Juhasz et al. (2005) presented both normally unspaced and spaced English compound words in sentences and recorded readers’ eye movements as they read these sentences. In addition, compounds were either shown with their correct spatial layout or their incorrect spatial layout. Inserting a space into a normally unspaced compound word severely hampered reading as indicated by much longer gaze durations on the compound. However, when the space was removed from a normally spaced compound word, there was no significant disruption in reading. Juhasz et al. (2005) concluded that there are two possible reasons for the reading advantage observed for compounds presented without interword spaces. One is that the unspaced format may make access to the whole-word representation easier and facilitate meaning assignment for the compound. The second is related to issues of visual acuity and the eye movement system. Parafoveal word length provides a strong cue for where to target the eyes in an upcoming word (Inhoff et al., 2003; Juhasz et al., 2008; White et al., 2005a). Readers typically land between the beginning and middle of words (the preferred viewing location, Rayner, 1979). This strategy might cause readers to land somewhat further into the unspaced compound than into the spaced compound. As a result the eyes would be closer to the 2nd lexeme, allowing more efficient information uptake because the letters of the 2nd lexeme would be in an area with higher visual acuity.

The boundary paradigm has usually been used to investigate parafoveal processing of words not yet fixated. However, Hyönä et al. (2004; Pollatsek & Hyönä, 2005) recently used this technique to examine parafoveal processing of the end of a long Finnish compound while the beginning of the compound was fixated. Finnish compounds were presented in sentences and

the boundary was set at the end of the first lexeme in the compounds. Prior to fixating the second lexeme, a parafoveal preview was provided consisting either of (1) the entire lexeme (full preview) or (2) only the first two letters (with the remaining letters replaced with visually similar letters). Having only the partial preview of the second lexeme increased reading time on that lexeme (a measure referred to as *subgaze2* which also includes regressions back to the first lexeme) by 70 ms. In contrast, the parafoveal preview condition did not influence reading times on the first lexeme prior to crossing the boundary. Thus, the lexemes appeared to be identified sequentially during reading, much as two separate words would be.

One interesting point from the Hyönä et al (2004) study was the size of the parafoveal preview effect: 80 ms. In a meta-analysis of prior parafoveal preview studies, Hyönä et al. reported that the average effect size was less than 14 ms for experiments where a partial parafoveal preview was provided in which the first 2-3 letters were identical to those in the target word. Thus, the 80 ms effect in the Hyönä et al. study was much larger than would be expected on the basis of prior experiments. There are two possibilities for the larger-than-expected effect of parafoveal preview for the second lexeme. The first possibility is related to low-level oculomotor factors. Readers' eyes were most likely closer to the boundary in the unspaced compound word than in past parafoveal preview experiments in which between-word manipulations were used. Therefore, the incorrect letter information from the partial parafoveal preview would have been closer to the fixation point in these words. However, there may be another explanation as well. Hyönä et al. suggested that the second lexeme may capture attention earlier than would be expected in experiments in which parafoveal preview is manipulated between words. Thus the large preview benefit may have been due to the fact that the second lexeme is part of a larger linguistic unit.

It is possible that both of these factors – (a) the landing position being closer to the boundary change for unspaced compounds and (b) the fact that compounds represent a larger linguistic unit – contributed to the large preview effect for the second lexeme in Hyönä et al. (2004). The English language allows a perfect test of the explanatory value of these two factors. Unspaced and spaced compounds can both be thought of as single linguistic units, as described above. Yet the inter-lexeme space allows a test of the oculomotor explanation of the large preview benefit effect discovered by Hyönä et al. (2004). In the current Experiment 1, both unspaced and spaced English compound words were used and a parafoveal preview manipulation similar to that used by Hyönä et al was employed. Experiment 2 tested whether there was a difference in parafoveal processing between spaced compounds and non-compound adjective-noun pairs. The results should inform our understanding of the processing of parafoveal information within and between words.

Experiment 1

Method

Participants. Twenty-eight participants from the University of Massachusetts, Amherst community participated either for course credit or for payment (\$11 per hour). All participants had normal uncorrected vision or wore contact lenses and were native speakers of English.

Apparatus. A Fourward Technologies Dual Purkinje eye-tracker (Generation 5) recorded participants' eye movements. In order to minimize head movements, a bite-bar was prepared and head rests were used. Participants read binocularly, although eye movements were recorded only from the right eye. The eye-tracker has a resolution of less than 10 minutes of arc. Participants were seated 61 cm from the computer screen where sentences were presented. At this distance,

3.8 characters subtend 1 degree of visual angle. The sentences were displayed on a 15-inch NEC MultiSync 4FG monitor in a fixed-width font and eye movements were recorded by a PC.

Materials. Twenty-four unspaced compound words were selected. These unspaced compounds ranged from 10 to 13 characters ($M=10.9$, $SD=0.88$). The average whole word frequency for the compounds was 1.78 per million (Educator's Word Frequency Guide; Zeno, Ivens, Hillard, & Duvvuri, 1995). First lexemes ranged from 6 letters to 8 letters ($M=6.3$, $SD=0.64$) with an average frequency of 71 per million. Second lexemes ranged from 4 letters to 6 letters ($M=4.6$, $SD=0.58$) with an average frequency of 181 per million. These unspaced compounds were matched to 24 spaced compound words which ranged from 11 to 14 characters ($M=11.96$, $SD=0.86$, with the space counting as a character). Whole word frequencies do not exist for spaced compounds. The first lexemes ranged from 6 to 8 letters ($M=6.4$, $SD=0.72$) with an average frequency of 54 per million. Second lexemes ranged from 4 to 6 letters ($M=4.5$, $SD=0.59$) with an average frequency of 167 per million. Lexeme lengths and frequencies did not differ significantly between unspaced and spaced compounds (all $t_s < 1$).

Sentence frames were created for each pair of unspaced and spaced compounds which were the same through the post-target word and then formed a meaningful completion after. Second lexeme constraint was assessed through a cloze norm. Eighteen Wesleyan University participants, who did not participate in the eye tracking portion of the study, were presented with the sentences through the first lexeme. They were also made aware of whether the second lexeme should be part of an unspaced compound or part of a spaced compound expression. Nine participants judged each version of the sentences. Second lexeme constraint did not vary as a function of compound type (Unspaced = 40% completion, Spaced = 39% completion, $t < 1$).

Two parafoveal previews of each second lexeme were prepared. The full preview consisted of all letters of the second lexeme, whereas the partial preview preserved the first two letters of the second lexeme but the remaining letters were replaced with visually similar letters. Parafoveal preview was manipulated using the boundary paradigm (Rayner, 1975). The invisible boundary was set immediately after the last letter of the first lexeme. When the readers' eyes crossed over this boundary, the preview was replaced by the intact second lexeme. Sentences are displayed in Appendix A.

Procedure: On arrival to the experiment, participants read and signed an informed consent form which also detailed the instructions for the experiment. Then a bite bar was prepared for each participant. A nine-point horizontal calibration routine was performed after the eye tracker was adjusted. The accuracy of the calibration was checked after each sentence by the experimenter and another calibration was performed whenever necessary. Participants read at their own rate and pressed a button when they finished reading each sentence. Comprehension was checked on approximately 25% of the sentences during the experiment by presenting the participant with a comprehension question that could be answered by the press of a button. Participants were over 96% accurate on the comprehension questions. In total, participants read 114 sentences: 24 experimental sentences were randomly intermingled with 85 filler sentences preceded by five practice sentences.

Design and Data Analysis: The data were analyzed using 2 x 2 Analyses of Variance (ANOVA). The first factor was the type of compound word (unspaced or spaced) and the second was the parafoveal preview of the second lexeme (full or partial). Error variance was computed over participants (*F1*) and items (*F2*). Both factors were considered within-participants, as all participants viewed all conditions. Both factors were also considered within-

items since the two types of words were matched on sentence context through the post-target word. The order of presentation of the sentences was randomized. The conditions were counterbalanced such that each participant viewed an equal number of items in each experimental condition.

Results

Trials were removed from the analysis if there was a blink or track loss of any kind on the pre-target, target, or post-target word. In addition, trials were removed if the display change did not occur at the correct time. In all, these criteria led to the removal of approximately 11.6% of the data. In addition, fixation durations that were 2.5 standard deviations above the condition mean were removed from each dependent measure separately (average data removed = 2.11%, range 1.35% to 3.20%). In addition, fixations shorter than 80 ms and on adjacent characters were combined. Fixations shorter than 100 ms or longer than 1000ms were eliminated by the data analysis software.

In order to examine the time-course of processing, a number of dependent measures were analyzed. Standard measurements of word recognition in reading will be reported. *First fixation duration* is the duration of the first fixation on a region independent of how many fixations the region ultimately receives. *Gaze duration* is the sum of fixations on a region before the readers move their eyes to a different region. We will also examine a measure referred to as *subgaze2*. This is the time spent fixating on the second lexeme including any regressions back to the prior lexeme before moving off of the compound expression. We will also discuss the percent of trials that the second lexeme was fixated and the percent of trials on which there was a regression back from the second lexeme. Since we are primarily concerned with the effect of denying the reader the final letters – while preserving the first two - in the preview on their encoding of the second

lexeme, we will begin with analyses of fixations in that region. We will then explore possible effects of the display change on fixations on the first lexeme.

Second Lexeme Measures. Condition means for the second lexeme are presented in Table 1. Because we are most interested in how long it took to process the second lexeme from when it was first fixated until the reader went on to the next word, we view the *subgaze2* as the most informative measure; however, we also report (and discuss first) the more standard first fixation duration and gaze duration measures.

Insert Table 1 about here

First fixations on the second lexeme were 36 ms longer when a partial parafoveal preview was presented as opposed to a full preview, $F(1,27) = 16.96$, $MSe = 2181$, $p < .001$; $F(1,22) = 16.29$, $MSe = 2131$, $p < .01$. In addition, there was an interaction between type of compound and parafoveal preview that was marginally significant by participants, $F(1,27) = 3.88$, $MSe = 2359$, $p < .06$) and fully significant by items ($F(1,22) = 4.31$, $MSe = 2434$, $p < .05$). The interaction stems from the fact that the effect of partial parafoveal preview was large (54ms) and highly significant for unspaced compounds, $t(27) = 3.84$, $p = .001$; $t(22) = 3.73$, $p = .003$, whereas it was small for spaced compounds (19 ms) and at best marginally significant, $t(27) = 1.64$, $p = .112$; $t(22) = 1.98$, $p = .061$. There was no main effect of type of compound ($F_s < 1$).

Gaze durations on the second lexeme were 44 ms longer when a partial parafoveal preview was provided, $F(1,26) = 18.50$, $MSe = 2894$, $p < .001$; $F(1,22) = 28.24$, $MSe = 1635$, $p < .001$. However, while the numerical pattern was similar to first fixation, in that there was a larger effect of parafoveal preview for unspaced compounds (57 ms) than for spaced compounds (31 ms), the interaction was not significant ($p_s > .10$). There was again no main effect of compound type ($F_s < 1$).

Subgaze2 reflects time spent reading the second lexeme plus any regressions to the first lexeme and subsequent re-readings of the second lexeme. Again, there was a highly significant 83 ms main effect of parafoveal preview in this measure, $F(1,26) = 28.43$, $MSe = 6413$, $p < .001$; $F(1,22) = 38.26$, $MSe = 4783$, $p < .001$, and a significant interaction, $F(1,26) = 4.96$, $MSe = 3858$, $p < .05$; $F(1,22) = 7.86$, $MSe = 3663$, $p < .05$, indicating a larger effect of parafoveal preview for unspaced compounds (109 ms) than for spaced compounds (56 ms). However, unlike with the first fixation duration, both effects were significant (unspaced: $t(26) = 4.75$, $p < .001$; $t(22) = 5.16$, $p < .001$; spaced: $t(26) = 3.61$, $p = .001$; $t(22) = 4.38$, $p < .001$). There was also a marginally significant effect of type of compound by participants, $F(1,26) = 2.97$, $MSe = 3657$, $p < .1$, but not by items ($p > .1$).

We also examined two fixation probability measures for the second lexeme: The percentage of trials when it was fixated during first pass, and the percentage of times a regression was made back from the second lexeme to an earlier region in the sentence. Both of these measures showed a main effect of compound type. More regressions (7%) were made from an unspaced compound's second lexeme, $F(1,27) = 4.24$, $MSe = 338$, $p < .05$; $F(1,23) = 7.92$, $MSe = 220$, $p = .01$, but the second lexeme in a spaced compound was 22% more likely to be fixated, $F(1,27) = 27.48$, $MSe = 491$, $p < .001$; $F(1,23) = 36.07$, $MSe = 317$, $p < .001$. A partial parafoveal preview resulted both in more regressions out of the second lexeme, $F(1,27) = 15.89$, $MSe = 178$, $p < .001$; $F(1,23) = 12.95$, $MSe = 234$, $p < .01$, and a greater probability of fixating the second lexeme, $F(1,27) = 10.38$, $MSe = 447$, $p < .01$; $F(1,23) = 9.73$, $MSe = 259$, $p < .01$. These two factors did not interact in either dependent measure ($ps > .1$).

First Lexeme Measures: Fixation durations on the first lexeme, prior to the display change, were also analyzed to examine whether processing of the second lexeme has any effect

on eye movements prior to when it is fixated or whether the lexemes have sequential control over eye movements. Condition means for the first lexeme are also presented in Table 1.

First fixations on the first lexeme were 26 ms longer when the compound was unspaced compared to spaced, $F(1,26) = 17.20$, $MSe = 1042$, $p < .001$; $F(1,23) = 17.19$, $MSe = 1046$, $p < .001$. However, parafoveal preview of the second lexeme did not significantly influence first fixations on the first lexeme ($F_s < 1$), nor was there an interaction between factors ($F_s < 1$). Similarly, gaze durations on the first lexeme were 43 ms longer on unspaced compounds than on spaced compounds, $F(1,26) = 17.43$, $MSe = 2920$, $p < .001$; $F(1,23) = 25.12$, $MSe = 2205$, $p < .001$. Although there was also a marginal 15 ms effect of parafoveal preview in the participant analysis, $F(1,26) = 3.94$, $MSe = 1645$, $p < .1$, indicating that preview of the second lexeme may have influenced time spent processing the first lexeme, this effect was not close to significant by items ($F < 1$) and did not interact with type of compound in either analysis ($F_s < 1$) suggesting it is spurious. However, it is possible that the inclusion of long first lexemes (7-8 letters) may have masked the effect of second lexeme preview on first lexeme processing. As mentioned in the introduction, models such as SWIFT (Engbert et al., 2005) posit that parallel lexical processing can only occur within a gradient of attention consisting of the perceptual span. Thus, parafoveal processing of the second lexeme would be expected to influence processing of the first lexeme only when both the first and second lexeme are within the perceptual span. In order to provide the strongest test of a parafoveal-on-foveal effect, a post-hoc analysis was conducted using only items where the first lexeme was 6 letters in length. Participant analyses were conducted on first fixation and gaze duration on the first lexeme. Consistent with the initial analyses, there were main effects of compound type in both dependent measures (First Fixation $F(1,26) = 22.40$,

MSe = 1237, $p < .001$; Gaze Duration $F(1,26) = 17.99$, MSe = 3271, $p < .001$) but no main effect of parafoveal preview ($ps > .1$) and no interaction ($Fs < 1$).

Whole Compound Measures: To examine a low-level oculomotor explanation for the interaction between type of compound and parafoveal preview reported above, the position of the first fixation on the whole compound was analyzed. Readers' eyes landed, on average, 3.31 characters into the compounds. However, there were no significant effects in this analysis (all $Fs < 1$).

The first-pass duration on the whole compound region was also analyzed (also presented in Table 1). This measure sums all fixations on either the first and second lexeme prior to the readers' eyes moving to an earlier or later region of the sentence. This measure thus reflects time spent on the first lexeme prior to the display change, time spent on the second lexeme after the display change, regressions back to the first lexeme after fixating the second lexeme and any subsequent rereading of the second lexeme. There was a significant effect of partial parafoveal preview. Compounds with a partial preview received 103 ms longer first-pass durations, $F(1,27) = 27.02$, MSe = 10926, $p < .001$; $F(1,23) = 32.78$, MSe = 6359, $p < .001$. However, even though the preview effect was 125 ms for unspaced compounds and only 81 ms for spaced compounds, neither the interaction between type and parafoveal preview nor the main effect of compound type were significant ($ps > .1$).

Additional Analyses: In order to further explore the interaction between compound type and preview reported in the second lexeme analyses, items were categorized according to the frequency of the first lexeme. A median frequency split was conducted to isolate the 24 compounds with the lowest frequency first lexeme and the 24 compounds with the highest frequency first lexeme. The low frequency first lexemes had frequencies ranging from 5 to 44

per million with a mean of 22.8. Of these, 13 were spaced compounds and 11 were unspaced compounds. The high frequency first lexemes had frequencies ranging from 45 to 359 per million with a mean of 103.7. Eleven of these were spaced compounds and 13 were unspaced compounds. Due to the small number of items in each condition, only participant analyses were conducted and this was done separately for the low frequency first lexeme compounds and the high frequency first lexeme compounds. Only participants with observations in each of the four cells contributed to the analyses. For the compounds with a low frequency first lexeme, the main effect of parafoveal preview effect was significant on several measures (First fixation on Lexeme 2, $F(1,19) = 9.01$, $MSe = 2817$, $p < .01$; Gaze duration on Lexeme 2, $F(1,19) = 8.88$, $MSe = 3460$, $p < .01$; Subgaze 2, $F(1,17) = 11.50$, $MSe = 17631$, $p < .01$). For these compounds, there were no significant effects of compound type ($ps > .1$) and no interaction between compound type and preview ($ps > .1$). The results for compounds with a high frequency first lexeme mirrored the overall analyses in demonstrating both main effects of preview (First fixation on Lexeme 2, $F(1,16) = 5.26$, $MSe = 4420$, $p < .05$; Gaze duration on Lexeme 2, $F(1,15) = 11.11$, $MSe = 3831$, $p < .01$; Subgaze 2, $F(1,17) = 31.20$, $MSe = 5870$, $p < .001$) and an interaction between compound type and preview (First fixation on Lexeme 2, $F(1,16) = 5.61$, $MSe = 3482$, $p < .05$; Gaze duration on Lexeme 2, $F(1,15) = 5.34$, $MSe = 3718$, $p < .05$; Subgaze 2, $F(1,16) = 4.11$, $MSe = 8659$, $p < .06$). However, for compounds with high frequency first lexemes there was also a main effect of compound type on the second lexeme (First fixation on Lexeme 2, $F(1,16) = 4.66$, $MSe = 1835$, $p < .05$; Gaze duration on Lexeme 2, $F(1,15) = 4.83$, $MSe = 2029$, $p < .05$; Subgaze 2, $F(1,16) = 5.16$, $MSe = 4940$, $p < .05$) indicating processing of the second lexeme is more difficult for unspaced compounds than for spaced compounds.

Discussion

Experiment 1 produced a clear pattern of results. Changing all but the first two letters of the second lexeme before it was fixated disrupted processing considerably for both unspaced and spaced compounds. This effect was significantly larger for unspaced compounds than spaced compounds as indicated by the interaction of parafoveal preview by compound type in first fixation on the second lexeme and subgaze2. This effect was qualified by the additional analyses which suggested that the interaction was only significant for compounds with a high-frequency beginning lexeme. Foveal load (i.e. how difficult it is to process the currently fixated word) has been found previously to modulate parafoveal preview effects (e.g. Henderson & Ferreira, 1990; White, Rayner, & Liversedge, 2005b) and is an important aspect of multiple models of eye movements in reading. These models suggest that when the currently fixated word is easy to process, more preview of the upcoming word is extracted from the parafovea. It is therefore not surprising that the interaction between compound type and second lexeme preview would be more robust for compound words with a high frequency first lexeme. While an interaction was not observed on first-pass duration on the entire compound, this is a relatively insensitive measure since it takes into consideration time spent fixating the first lexeme prior to the display change. The first lexeme analyses did not produce any reliable effects of second lexeme parafoveal preview - although there were effects of compound type such that unspaced compounds received longer processing on the first lexeme. The finding that unspaced compounds have longer initial fixations is consistent with Juhasz et al. (2005) who also examined processing on unspaced and spaced compound words.

The finding that the partial preview effect for the second lexeme was larger for the unspaced compounds has two plausible causes (not mutually exclusive). First, given that the initial landing position on the first lexeme is about the same for the spaced and unspaced

compounds, the letter information from the second lexeme is one space farther from fixation and thus of poorer visual quality. Moreover, any benefit in processing the second lexeme from the space due to reducing lateral inhibition would be irrelevant to the preview manipulation as that benefit would be likely only for the first letter or two of the second lexeme – and those letters are present in both preview conditions. Second, however, the increased benefit for the unspaced compounds could have a “deeper” cause: readers may be significantly more likely to treat the unspaced compound as a linguistic unit and thus be more likely to spread their focus of spatial attention over the whole compound. We will discuss these alternatives more fully in the General Discussion. However, the data indicate that the effect observed even for spaced compounds was larger than one would predict on the basis of past boundary experiments. As noted earlier, in their meta-analysis of past parafoveal preview experiments, Hyönä et al. (2004) reported an average effect size of +14ms when a partial parafoveal preview is provided when the first 2-3 letters are kept identical and the rest are visually dissimilar and an effect of -7ms when the first 2-3 letters are identical and the rest are visually similar to the target (this is the condition used in the present study). Yet, in the present experiment we observed a 31 ms effect on gaze duration in the spaced compound condition. Thus, the data suggest that even the spaced compounds, which are familiar to speakers of English, are processed as a linguistic unit at least some of the time. Thus, either attention may shift to the second lexeme of the spaced compound earlier than when two successive words from non-lexicalized expressions are encountered in text; or on some fraction of the trials, the spaced compound is an attentional unit (and processed in parallel). In order to test whether spaced compounds are processed differently than other two word sequences, Experiment 2 was conducted, which used the same spaced compounds and sentence frames as Experiment 1, but each spaced compound was instead matched to an adjective-noun

pair containing the same second noun. This allowed us to compare the size of the parafoveal preview effect on the same noun when it was part of a familiar spaced compound to when it was part of an adjective-noun pair that is not a part of a readers' mental lexicon. If the large partial preview benefit observed in Experiment 1 was due to the fact that the spaced compounds are part of the larger linguistic unit, then we should observe a smaller effect on the noun for the adjective-noun pairs.

Experiment 2

Method

Participants. Thirty-six University of Massachusetts, Amherst community members participated for the same compensation; they met the same criteria as in Experiment 1.

Apparatus. The apparatus was identical to Experiment 1.

Materials. The same 24 spaced compound words and corresponding sentence frames from Experiment 1 were used. In addition, an adjective was selected which fit with the second lexeme of each spaced compound, but did not form an existing spaced adjective-noun compound (as assessed through an online dictionary). These adjectives were matched as closely as possible to the nouns of the spaced compounds in length ($M = 6.3$, $SD = .91$) and frequency ($M=49.9$, $SD=51.3$) and did not differ significantly from the first lexeme of the spaced compounds on either measure (both $t_s < 1$). If possible, sentences were kept the same for both the spaced compound and the adjective-noun pair. For seven pairs, sentences needed to be slightly reworded to accommodate the adjective-noun pair. In order to assess the predictability of the noun, a cloze task was conducted with 16 Wesleyan University undergraduates who did not take part in the norming for Experiment 1 or the eye-tracking studies. Sentences were provided up through the first lexeme or the adjective and participants were asked to provide a word that could fit as the

next word in the sentence. Unlike in Experiment 1, the cloze task for Experiment 2 did not indicate to participants that the noun provided should be part of a spaced compound. (However, logically, it needs to be so in the spaced compound condition.) Eight participants provided cloze completions for each version of the paired sentences. Not surprisingly (because the constraint is greater in the compound noun condition), the target noun was provided as a completion significantly more in the spaced compound sentences (27%) than the adjective-noun sentences (6.8%; $t(23)=3.46$, $p<.01$). The two parafoveal preview conditions used were the same as in Experiment 1, with the boundary immediately after the first lexeme of the spaced compound and adjective. The stimuli are presented in Appendix B.

Procedure: The procedure was virtually identical to that in Experiment 1. In total, participants read 104 sentences: 24 experimental sentences mixed randomly with 76 filler sentences preceded by five practice sentences. Participants were over 95% accurate on the comprehension questions.

Design and Data Analysis: The data were analyzed similarly to those in Experiment 1.

Results

Trials were removed using the same criteria as for Experiment 1 (13.4% of the data). In addition, fixation durations that were 2.5 standard deviations above the condition mean were removed from each dependent measure separately (average data removed from all measures= 2.42%, range 1.74% to 2.81%). In addition, fixations shorter than 80 ms and on adjacent characters were combined. Fixations shorter than 100 ms or longer than 1000 ms were eliminated by the data analysis software. The same set of dependent measures is reported as for Experiment 1.

Second Lexeme Measures: Reading times on the second lexeme are critical for assessing the role of parafoveal preview in spaced compounds as opposed to non-lexicalized adjective-noun pairs. Condition means for the second lexeme are presented in Table 2. First fixations on the second lexeme were 29 ms longer when a partial preview was presented in the parafovea ($F(1,35) = 27.94$, $MSe = 1106$, $p < .001$; $F(1,22) = 16.29$, $MSe = 2131$, $p < .01$). On average, first fixations on the noun of an adjective-noun pair were 15 ms longer than on the second lexeme of a compound word; however, although this effect was significant by participants ($F(1,35) = 9.78$, $MSe = 755$, $p < .01$), it was not by items ($p > .1$). In addition, while the effect of parafoveal preview was numerically larger for the spaced compounds (35 ms) than for the adjective-noun pairs (23 ms), the interaction between variables was not close to significant ($ps > .1$). Similarly, gaze durations were 13 ms longer for adjective-noun pairs than for spaced compounds, but the effect was only significant over participants, $F(1,35) = 15.65$, $MSe = 1721$, $p < .001$; $F(1,23) = 1.55$, $MSe = 999$, $p > .2$, and were 27 ms longer with a partial preview of the second lexeme than with a full preview, $F(1,35) = 15.65$, $MSe = 1721$, $p < .001$; $F(1,22) = 28.24$, $MSe = 1635$, $p < .001$. Similar to first fixation duration, the numerical trend was for a larger parafoveal preview effect for the spaced compounds (34 ms) than for the adjective-noun pairs (20 ms), but the interaction was not significant ($ps > .2$).¹

Insert Table 2 about here

The pattern of data was similar for subgaze2. The 39 ms main effect of parafoveal preview was significant, $F(1,35) = 16.71$, $MSe = 3322$, $p < .001$; $F(1,23) = 22.76$, $MSe = 1866$,

¹ Similar to Experiment 1, second lexeme durations were analyzed separately for low frequency and higher frequency first words (according to a median split on the first word frequency). No significant interactions between item type and second lexeme preview were observed. All viewing duration measures indicated an effect of parafoveal preview, although this effect failed to reach significance in subgaze2 and gaze duration for items with a low frequency first word ($ps > .2$). There were also marginal effects of item type on subgaze 2 and first fixation on lexeme 2, mirroring the pattern from the overall analysis.

$p < .001$), and the nouns in adjective-noun pairs received 24 ms longer subgaze2 durations than the second lexeme in the spaced compounds, $F(1,35) = 9.86$, $MSe = 2077$, $p < .01$; $F(1,23) = 5.86$, $MSe = 1618$, $p < .05$. The interaction between factors was marginally significant by participants, $F(1,35) = 2.99$, $MSe = 1530$, $p < .1$; $F < 1$, and follow-up t-tests were conducted. The 50 ms parafoveal preview effect for spaced compounds was significant, $t(35) = 4.84$, $p < .001$; $t(23) = 3.90$, $p < .01$, but the 28 ms parafoveal preview effect for adjective-noun pairs just missed significance in the item analysis, $t(35) = 2.21$, $p < .01$; $t(23) = 2.01$, $p < .1$.

As in Experiment 1 we examined the probability of fixating the second lexeme and the percentage of regressions back from the second lexeme. Unlike Experiment 1, there were no significant effects on regressions back from the second lexeme, although the effect of parafoveal preview was marginal by items ($F(1,35) = 1.46$, $MSe = 164$, $p > .2$; $F(1,23) = 3.82$, $MSe = 34$, $p < .1$). The 6.0% main effect of parafoveal preview on the percentage of fixations on the second lexeme was significant, $F(1,35) = 6.33$, $MSe = 210$, $p < .05$; $F(1,23) = 7.16$, $MSe = 117$, $p < .05$, but there were no effects of type of word pair and no interaction ($F_s < 1$).

First Lexeme Measures: First fixation durations and gaze durations on the first lexeme of the spaced compound and the adjective were also analyzed to examine whether there were any effects of parafoveal preview prior to fixation on the second lexeme. Condition means for the first lexeme are presented in Table 2. These analyses revealed no significant difference in processing between adjectives and nouns ($p_s > .1$) and no main effects of parafoveal preview of the second lexeme ($p_s > .2$). There was a surprising interaction between variables that was not significant by participants ($p_s > .1$) but was marginally significant for first fixation duration by items, $F(1,23) = 3.93$, $MSe = 383$, $p < .1$, and significant in gaze duration by items, $F(1,23) = 5.30$, $MSe = 946$, $p < .05$. The interaction is due to a cross-over pattern in the means whereby the

difference between full and partial preview was negative for spaced compounds (FFD -9 ms; GD -8 ms) but positive for adjective-noun pairs (FFD 7ms; GD 20 ms). However, analysis of the simple effects indicate that none of these effects are, in fact, significant (all p s $>.05$) and thus there is no evidence that parafoveal preview of the second lexeme was affecting processing of the first lexeme or of the adjective. Similar to Experiment 1, we also conducted by-participant analyses for the shortest first lexemes (5-6 letters) separately to provide the strongest test of parafoveal-on-foveal effects. There were no significant effects for first fixation duration or gaze duration for these items (all F s <1).

Whole Region Measures: First pass time on the whole compound region was also analyzed. There was a significant 50 ms effect of partial parafoveal preview on first-pass duration, $F(1,35) = 11.96$, $MSe = 7682$, $p <.01$; $F(1,23) = 10.42$, $MSe = 5027$, $p <.05$, and a 31 ms effect of item type, with adjective noun pairs having longer first-pass times, $F(1,35) = 9.88$, $MSe = 3425$, $p <.01$; $F(1,23) = 3.76$, $MSe = 7469$, $p <.1$. Surprisingly, the preview effect was actually larger for the adjective-noun pairs than for the spaced compounds (78 ms vs. 22 ms). However, the interaction was significant only in the participant analysis ($F(1,35) = 6.61$, $MSe = 4406$, $p <.05$; $F(1,23) = 3.38$, $MSe = 6362$, $p <.1$).

Discussion

The size of the partial parafoveal preview manipulation for eye movement measures on the second lexeme of spaced compounds was very consistent between Experiment 1 (31 ms Gaze, 56 ms Subgaze2) and Experiment 2 (34 ms Gaze, 50 ms Subgaze2). However, though this effect was numerically larger for spaced noun-noun compounds than for the adjective-noun word pairs (20 ms Gaze, 28 ms Subgaze 2), none of the interactions between item type and parafoveal preview on the second lexeme reached significance. Thus, we can only conclude that parafoveal

processing of the second noun in the spaced noun-noun compounds and non-lexicalized adjective-noun word pairs is roughly equivalent.

There was evidence that processing of the non-lexicalized adjective-noun word pairs takes more time than that of the lexicalized spaced compounds. Adjective-noun pairs had significantly longer subgaze² durations and also longer overall first-pass durations. This could be due to the fact that no whole-word representation exists for these word pairs and their meaning must therefore be computed online as opposed to accessed from the readers' mental dictionaries. Noun processing may also take more time for the adjective-noun pairs because the nouns are less predictable following the adjectives than following the first noun in the spaced compounds (as indexed by the cloze rating). First-pass durations also suggested that overall processing of the adjective-noun pairs was hurt more by the partial parafoveal preview of the noun than with the spaced noun-noun compounds. Since the noun is the semantic head of these expressions, this again may indicate that the meaning computation process was more difficult for the adjective-noun expressions, especially when a parafoveal preview of the second lexeme was denied.

What is perplexing about this experiment is that the size of the decrement due to the changed final letters of the parafoveal preview for the adjective-noun pairs was still larger than one would have expected on the basis of past experiments. (Besides the preview preserving the first two letters, the changed letters were kept visually similar to those in the target word.). That is, according to the meta-analysis conducted by Hyönä et al. (2004), past studies would predict no effect of partial preview in this case (the average is -7 ms), However, we observed a robust effect of 20 ms in gaze duration on the noun, which is thus closer in size to experiments which have kept the first two to three letters the same and replaced the remaining letters with *xs*.

Reasons for the discrepancy between these experiments and past experiments will be addressed in the General Discussion.

General Discussion

The purpose of the experiments reported here was to examine parafoveal processing on the fixated word. To date, most experiments examining parafoveal processing in reading have examined parafoveal processing to the right of the fixated word (for text running left to right). An exception to this are the studies by Hyönä et al (2004) and Pollatsek & Hyönä (2005) that examined parafoveal processing of the second lexeme in long Finnish compound words. Hyönä et al found that when a partial preview of the second lexeme was provided, subgaze2 duration (a measure of processing on the compound after the display change has occurred) was 80 ms longer than when a full preview of the second lexeme was provided. In contrast, the incorrect information from the partial second lexeme preview did not affect reading time on the first lexeme prior to the display change. Thus, Hyönä et al. concluded that lexemes are processed sequentially in Finnish compounds. The results of Experiment 1 suggest that lexemes in English unspaced compound words are also processed sequentially. Having a partial preview of the second lexeme did not significantly influence time spent on the first lexeme prior to the display change (i.e. there was no evidence of parafoveal-on-foveal effects). However, there were some non-significant trends for parafoveal-on-foveal effects for spaced compounds and adjective-noun pairs.

These parafoveal-on-foveal effects are sometimes reported in the eye-movement literature (see e.g., Kennedy & Pynte, 2005) and constitute one of the main pieces of evidence for parallel lexical processing. If two words are processed at the same time in reading, one would expect that characteristics of BOTH words should influence fixation durations. Thus, having an

inaccurate preview of a parafoveal word should influence fixation durations on the currently fixated word. It is often hard to argue against parafoveal-on-foveal effects, because this amounts to arguing for a null effect. The best way to check for parafoveal-on-foveal effects is to conduct multiple analyses on the conditions where parafoveal-on-foveal effects should be found. In Experiments 1 and 2 supplementary analyses were reported on the items containing the shortest first lexemes. No significant parafoveal-on-foveal effects were observed. As a final test of parallel lexical processing, an additional ANOVA was conducted pooling the data from the two experiments. Since sentences were almost identical for the two experiments, a by-items 3x2 ANOVA was conducted. The first factor included in the ANOVA was item type (unspaced compound, spaced compound, or adjective-noun pair) and the second factor was parafoveal preview of the second lexeme (full or partial). The dependent measures were first fixation duration and gaze duration on the first lexeme/adjective. While there were main effects of item type ($p < .001$), there were no significant effects of parafoveal preview and no significant interaction between variables ($p > .1$). The results of these combined analyses thus provide support for sequential attention shift models of eye movements during reading (such as E-Z Reader, Reichle et al., 1998; 2003). If parallel lexical processing occurs during reading, one would expect that what amounts to quite severe “mutilations” of the second lexeme of a compound word should have an effect on the eyes prior to it being fixated. The failure to find such evidence for parallel processing for these compounds calls into question the assumption that parallel lexical processing is the usual occurrence in reading.

Experiment 1 also confirmed the results of Hyönä et al (2004) by demonstrating a large effect of parafoveal preview of the second lexeme for the *subgaze2* measure for unspaced compounds. This effect (109ms) was larger than the effect observed for spaced noun-noun

compounds (56 ms). This result thus indicates that the finding of greater preview benefit for information from a lexeme within a word is not restricted to languages like Finnish where unspaced compounding is obligatory. The issue remains, however, as to why there is such a substantial difference between spaced and unspaced compounds.

As indicated earlier, one possibility is that the letter information in the second lexeme is, on average, about one character further from fixation for spaced compounds and thus the letter information from them is harder to extract. (Remember that the mean location of the first fixation on the first lexeme was the same for spaced and unspaced compounds.) Although this difference in location is likely to explain some difference between the two conditions, it seems unlikely to explain a 50 ms difference. That is, assuming a model such as E-Z Reader 9 (Pollatsek, Reichle, & Rayner, 2006), with the standard parameter values for decrease in letter-processing speed due to the eccentricity difference in the two conditions, one would predict at best a 10 ms difference in subgaze² even if one were contrasting a whole-word preview with a preview that preserved none of the letters. Obviously, the prediction for the difference between a whole-word preview and a partial preview would be even smaller². Thus, if eccentricity differences are an important part of the observed difference between the spaced and unspaced compound difference in Experiment 1, it must be that the eccentricity differences are interacting with the fact that the unspaced compound is an orthographic unit and the spaced compound is not.

² To be a bit more precise, a prediction of a 6 ms difference was generated contrasting a whole-word preview with a preview where none of the letters were preserved assuming (a) that the launch site of the saccade to the second lexeme was always at the average fixation location in Table 1 and (b) that the eyes would always land in the optimal location on the second lexeme. Adding in variability to assumption (a) and adjusting for the fact that the fixation location on the second lexeme is rarely optimal would each make the estimate a bit larger. We should also note that current versions of E-Z Reader do not have a complete enough model of word encoding to make predictions for partial preview conditions. (The values for the parameters were obtained from the corpus used to fit the model in Pollatsek et al., 2006.)

This raises the question, of course, as to what could be happening that is different when the compound is an orthographic unit and when it is not. However, as we argued above from the lack of preview effects on first-pass fixation measures on the first lexeme, there is no evidence that the two lexemes of the unspaced compounds are processed in parallel at the lexeme level. This would appear to leave two possibilities for why more preview information is extracted from the second lexeme when the compound is unspaced. One is that the shift of attention to the second lexeme is more rapid when the second lexeme is part of the same orthographic unit. This could mean that readers “cheat” a bit on when they decide the first lexeme has been processed when it is part of an unspaced compound and shift attention to the second lexeme before full processing of the first lexeme has completed. The other possibility is that, at levels below the lexical level, processing of the information from the two lexemes is in parallel. That is, processing of the letter information from all (or most) letters of the unspaced compound may occur in parallel. This parallel processing of orthographic information is consistent with serial attention shift models such as E-Z Reader (Pollatsek et al., 2006). In E-Z Reader, it is word identification that occurs sequentially; once a word is identified, attention shifts to the next word in text. Lower-level letter processing occurs in parallel in a pre-attentive visual processing stage. Among other things, this pre-attentive parallel processing of visual features is necessary to program a saccade to upcoming words in text. Indeed, positing that no processing of the letter information of the second lexeme occurs before the first lexeme has been identified would assume that the reader automatically knows where the lexeme boundary is before identifying the first lexeme. Thus, the greater preview benefit for unspaced compounds is plausibly because several of its letters have been fairly well identified before the parsing of the compound has progressed fully enough for spatial attention to narrow to focusing on the first lexeme. (This

would be before spatial attention shifts to the second lexeme to identify the lexeme as a whole.)

In either case, as noted above, the greater preview benefit due to a difference in attention is likely to be enhanced by the fact that the letter information is also closer to fixation.

Although the preview benefit in the unspaced compound condition was greater than in the spaced compound condition, the latter was larger than one would predict on the basis of past experiments where parafoveal preview was manipulated between words and a partial preview was provided (as reviewed in Hyönä et al., 2004). The effect on the gaze duration on the second lexeme was 31 ms in Experiment 1, whereas the review by Hyönä et al yielded an effect size of -7 ms across experiments with the same type of partial preview as was employed in the current experiments. One possibility for this larger effect in the current study was that the fact that, even though the spaced compound did not form an orthographic unit, it formed a linguistic unit that allowed attention to shift to the second lexeme earlier than is typically observed between words. This earlier attention shift would therefore produce a larger parafoveal preview of the second lexeme. However, a comparison between non-lexicalized adjective-noun word pairs and the same spaced compounds in Experiment 2 ruled out this linguistic unit hypothesis, as the adjective-noun pairs used in Experiment 2 were novel expressions that are not found in dictionaries. Although the adjectives were all plausible modifiers for the noun, the adjective-noun pairs were not lexicalized expressions and therefore their meaning must be computed online. However, analysis of gaze durations on the noun showed a relatively large effect of partial parafoveal preview (20 ms) following the adjectives – larger than one would have expected on the basis of past experiments. In addition, although this effect was numerically smaller than that observed with the spaced compounds (31 ms), the interaction between item type and preview did not reach significance.

One possible reason for the discrepancy in the size of the parafoveal preview effect for the current experimental items and those that are reported in the meta-analysis conducted by Hyönä et al. (2004) is differences between the items used. Five experiments were reported in the meta-analysis conducted by Hyönä et al. which most closely match the parafoveal preview conditions used in the present experiment. In the experiments that had conditions where the first 2-3 letters were kept intact and the remaining letters were replaced by visually dissimilar letters (Lima, 1987 Experiment 2; Henderson & Ferreira, 1990 Experiment 1 and 2; Inhoff, 1989 Experiment 3; Pollatsek et al. 1992), Hyönä et al report an average effect size of +14 ms. In the three experiments where the same manipulation as in the present experiment was used (Henderson & Ferreira, 1990 Experiment 1 and 2; Pollatsek et al 1992), an average effect size of -7 is reported. Examination of the items used in these experiments, however, show that a range of syntactic categories were employed consisting of verbs, adjectives, and nouns. These experiments therefore did not control the items to the same extent as in the current study.

Balota, Pollatsek, and Rayner (1985; see also Drieghe, Rayner, & Pollatsek, 2005) demonstrated that when a word is predictable from sentence context, parafoveal processing of that word is increased, leading to a larger parafoveal preview benefit than when the same word is in a neutral sentence context. In the present experiment the sentence frames were constructed so that (among other things) the target words were not very predictable. However, recent eye movement research has suggested that syntactic prediction does occur on-line (e.g. Staub & Clifton, 2006; Staub, Rayner, Pollatsek, Hyönä, & Majewski, 2007). These studies thus indicate that readers quickly compute the likelihood of the next word's syntactic category and this affects early fixation duration measures when that word is fixated. In the present Experiment 2, the syntactic category of the target word was almost always predictable. Thus this syntactic

predictability may have increased the parafoveal preview benefit on the noun, much as lexical predictability has been found to. Future research should directly examine the effects of syntactic predictability compared to lexical predictability on parafoveal processing.

In conclusion, the present experiments demonstrate that lexemes are processed sequentially in English unspaced compound words, converging with the results from Finnish compounds. They also demonstrate that there is a larger parafoveal preview benefit from the second lexeme for unspaced compounds than for spaced compounds. This larger benefit is likely due largely to differences in how attention is allocated for spaced and unspaced compounds, although differences in eccentricity of the letter information in the second lexeme may play a role as well. The syntactic predictability of the noun in spaced noun-noun compounds and adjective-noun word pairs may also allow more information to be extracted parafoveally than for typical words that are not predictable in any way and thus produce larger parafoveal preview benefits.

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.
- Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology*, *16*, 285-311.
- Ashby, J., Treiman, R., Kessler, B., & Rayner, K. (2006). Vowel processing during silent reading: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 416-424.
- Bai, X., Yan, G, Liversedge, S.P., Zang, C., & Rayner, K. (2008). Reading spaced and unspaced Chinese text: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, in press.
- Balota, D.A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, *17*, 364-390.
- Bertram, R., & Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory and Language*, *48*, 615-634.
- Bertram, R., & Hyönä, J. (2007). The interplay between parafoveal preview and morphological processing in reading. In R.G. van Gompel, M.H. Fischer, W.S. Murray & R.L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 391-407). Oxford: Elsevier Science.

- Chace, K.H., Rayner, K., & Well, A.D. (2005). Eye movements and phonological parafoveal preview: Effects of reading skill. *Canadian Journal of Experimental Psychology, 59*, 209-217.
- Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Eye movements and word skipped during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 954-969.
- Engbert, R., Nuthmann, A., Richter, E., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*, 777-813.
- Henderson, J.M., Dixon, P., Peterson, A., Twilley, L.C., & Ferreira, F. (1995). Evidence for the use of phonological representations during transsaccadic word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 82-97.
- Henderson, J.M. & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 417-429.
- Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eye-movement contingent display change study. *Memory & Cognition, 32*, 523-532.
- 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.
- Hyönä, J. & Pollatsek, A. (1998). The role of component morphemes on eye fixations when reading Finnish compound words. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1612-1627.

- Inhoff, A.W. (1989). Parafoveal processing of words and saccade computation during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 544-555.
- Inhoff, A.W., Radach, R., Eiter, B.M., & Juhasz, B. (2003). Distinct subsystems for the parafoveal processing of spatial and linguistic information during eye movements in reading. *Quarterly Journal of Experimental Psychology*, 56A, 803-827.
- Inhoff, A.W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, 40, 431-439.
- Juhasz, B.J. (2007). The influence of semantic transparency on eye movements during English compound word recognition. In R. vonGompel, W. Murray, and M. Fischer (Eds) *Eye movements: A window on mind and brain* (pp. 373-389). Boston: MA: Elsevier.
- Juhasz, B.J., Inhoff, A.W., & Rayner, K. (2005). The role of interword spaces in the processing of English compound words. *Language and Cognitive Processes*, 20, 291-316
- Juhasz, B.J., Starr, M. & Inhoff, A.W., & Placke, L. (2003). The effects of morphology on the processing of compound words: Evidence from naming, lexical decisions, and eye fixations. *British Journal of Psychology*, 94, 223-244.
- Juhasz, B.J., White, S.J., Liversedge, S.P., & Rayner, K. Eye movements and the use of parafoveal word length information in reading. *Journal of Experimental Psychology: Human Perception and Performance*, in press,
- Kambe, G. (2004). Parafoveal processing of prefixed words during eye fixations in reading: Evidence against morphological influences on parafoveal preprocessing. *Perception & Psychophysics*, 66, 279-292.
- Kennedy A, Pynte J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research*,

- 45, 153-68.
- Lima, S.D. (1987). Morphological analysis in sentence reading. *Journal of Memory and Language*, 26, 84-99.
- McConkie, G.W., & Rayner, K. (1975). The span of effective stimulus during a fixation in reading. *Perception & Psychophysics*, 17, 578-586.
- McConkie, G.W., & Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society*, 8, 365-368.
- Pollatsek, A., & Hyönä, J. (2005). The role of semantic transparency in the processing of Finnish compound words. *Language and Cognitive Processes*, 20, 261-290.
- Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 820-833.
- Pollatsek, A., Lesch, M., Morris, R.K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 148-162.
- Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology*, 52, 1-52.
- Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, 7, 65-81.
- Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception*, 8, 21-30.

- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372-422.
- Rayner, K., Balota, D.A., & Pollatsek, A. (1986). Against parafoveal semantic preprocessing during eye fixations in reading. *Canadian Journal of Psychology*, *40*, 473-483.
- 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.
- Rayner, K., McConkie, G.W., & Zola, D. (1980). Integrating information across eye fixations. *Cognitive Psychology*, *12*, 206-226.
- Rayner, K., & Pollatsek, A. (1989). *The Psychology of Reading*. NJ: Laurence Erlbaum Associates.
- Rayner, K., Well, A.D., & Pollatsek, A. (1980). Asymmetry of the effective visual field in reading. *Perception & Psychophysics*, *27*, 537-544.
- Reichle, E., Pollatsek, A., Fisher, D.L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*, 125-157.
- Reichle, E.D., Rayner, K., and Pollatsek, A. (2003). The E-Z Reader model of eye movement control in reading: Comparison to other models. *Behavioral and Brain Sciences*, *26*, 445-476.
- Schilling, H.E., Rayner, K., & Chumbley, J.I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition*, *26*, 1270-1281.

- Staub, A., & Clifton, C. (2006). Syntactic prediction in language comprehension: Evidence from either.....or. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 425-436.
- Staub, A., Rayner, K., Pollatsek, A., Hyönä, J., & Majewski, H. (2007). The time course of plausibility effects on eye movements in reading: Evidence from noun-noun compounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 1162-1169.
- White, S.J., Rayner, K., & Liversedge, S.P. (2005a). The influence of parafoveal word length and contextual constraint on fixation durations and word skipping in reading. *Psychonomic Bulletin & Review*, 12, 466-471.
- White, S.J., Rayner, K., & Liversedge, S.P. (2005b). Eye movements and the modulation of parafoveal processing by foveal processing difficulty: A reexamination. *Psychonomic Bulletin & Review*, 12, 891-896
- Zeno, S. M., Ivens, S. H., Hillard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Brewster, NJ: Touchstone Applied Science Associates.

Acknowledgments

This research was initiated when all of the authors were at the University of Massachusetts, Amherst. The research was supported by Grant HD26765 from the National Institute of Child Health and Human Development. Jukka Hyönä acknowledges the financial support of Suomen Akatemia (the Academy of Finland). Denis Drieghe is a postdoctoral fellow of the Fund for Scientific Research (Flanders, Belgium). We thank Martin Fischer and three reviewers for their helpful comments.

Table 1: First fixation duration and gaze durations on the first and second lexeme, subgaze2, number of regressions back from second lexeme, percentage of trials on which the second lexeme was fixated, first-pass duration and initial landing position on the entire region in Experiment 1.

	Unspaced Compound		Spaced Compound	
	Full Preview	Partial Preview	Full Preview	Partial Preview
First Fixation Duration on Lexeme 2	256 (66)	310 (46)	274 (46)	293 (43)
Gaze Duration on Lexeme 2	260 (66)	317 (57)	274 (42)	305 (57)
Subgaze2	265 (67)	374 (97)	272 (43)	327 (80)
Percent Regressions back from 2nd Lexeme	6.0 (15)	18.4 (27)	1.1 (4.2)	8.9 (15)
Percent of trials 2nd lexeme was fixated	53 (23)	69 (29)	78 (23)	88 (17)
First Fixation Duration on Lexeme 1	285 (52)	285 (33)	262 (38)	256 (40)
Gaze Duration on Lexeme 2	331 (71)	343 (70)	285 (50)	303 (72)
First-Pass Duration on Compound	436 (94)	561 (192)	480 (124)	561 (153)

Note: Numbers in parentheses are standard deviations. Fixation durations are in milliseconds.

Table 2: First fixation duration and Gaze durations on the first and second lexeme, subgaze2, number of regressions back from second lexeme, percentage of trials on which the second lexeme was fixated, and first-pass duration on the entire region in Experiment 2.

	Unspaced Compound		Spaced Compound	
	Full Preview	Partial Preview	Full Preview	Partial Preview
First Fixation Duration on Lexeme 2	253 (42)	288 (38)	273 (44)	297 (34)
Gaze Duration on Lexeme 2	265 (54)	299 (44)	286 (59)	306 (39)
Subgaze2	267 (54)	318 (48)	302 (63)	330 (55)
Percent Regressions back from 2nd Lexeme	6.0 (12)	9.5 (14)	7.4 (12)	9.1 (13)
Percent of trials 2nd lexeme was fixated	85 (18)	87 (17)	83 (24)	92 (12)
First Fixation Duration on Lexeme 1	264 (37)	260 (38)	263 (43)	270 (37)
Gaze Duration on Lexeme 2	286 (50)	284 (69)	283 (47)	303 (62)
First-Pass Duration on compound	500 (100)	522 (110)	503 (114)	581 (100)

Note: Numbers in parentheses are standard deviations. Fixation durations are in milliseconds.

Appendix A

Sentences used in Experiment 1. The unspaced compound is first shown, followed by the spaced compound (italicized). Participants only viewed one version of each sentence. The boundary was located immediately after the first lexeme in these compounds. Prior to fixating the second lexemes readers received either a correct preview or a partial preview (in parentheses).

1. Gary found an old basketball (badk)/ *tennis ball (badk)* in the corner of the garage.
2. The lights in the livingroom (rosw)/ *locker room (rosw)* were annoyingly bright.
3. Fred wanted a picture of the mountainside (silr)/ *mountain range (raoyz)* for his collection.
4. George didn't expect the pillowcase (cauv)/ *pillow fight (fiqbk)* in the attic to be so clean./ *so noisy.*
5. He wanted to buy a pocketknife (knzhw)/ *guitar string (stucap)* but he didn't have enough time.
6. The museum had a picture of a silversmith (smrbk)/ *dinner table (tahkm)* from the late 18th century.
7. The antique candlestick (stmel)/ *coffee table (tahkm)* in the dining room was very expensive.
8. John pulled out a piece of letterhead (hexb)/ *carbon paper (paqmo)* from his desk drawer.
9. They installed a new searchlight (liykb)/ *garage door (dosu)* next to the police station.
10. He didn't expect the thunderstorm (stcua)/ *traffic light (liykb)* to come up so suddenly.
11. Mary was pleased to see gingerbread (brczh)/ *health food (fomt)* in the new corner store.
12. The noise coming from the lumberyard (yaeb)/ *assembly line (lioc)* was surprisingly loud.
13. He wondered whether the congresswoman (wouxe)/ *exchange rate (rafx)* would speak/*change* before the end of the day.

14. The visitors to the former battlefield (fivtb)/ *prison camp (cavg)* were very moved.
15. For every policewoman (wouxe)/ *beauty queen (qucmo)*, morning exercises are a part of the daily routine.
16. Joan knew that one troublemaker (mahvu)/ *problem child (chrkf)* in the class could spoil it for everyone.
17. Mike needed to repair the clothesline (lioc)/ *kitchen sink (sirl)* before the week-end.
18. The experienced screenwriter (wrnkco)/ *travel agent (agmaf)* liked to keep a diary of the places he visited.
19. The people in the coffeehouse (hocaw)/ *concert hall (hatf)* relaxed and enjoyed the music.
20. The committee saw that the churchyard (yaeb)/ *tennis court (comvd)* badly needed some work.
21. Linda thought that her new screenplay (plcg)/ *credit card (caeh)* was an improvement over the old one.
22. Johnny wondered who left the cheesecake (cahv)/ *dollar bill (bifh)* on the kitchen counter.
23. A flash of light appeared in the thundercloud (clnrh)/ *mirror image (imnjw)* that Ralph was staring at.
24. Nancy was fumbling with the buttonhole (hodw)/ *rubber band (baof)* because it was so dark.

Appendix B

Sentences used in Experiment 2. . The spaced compound is first shown, followed by the adjective-noun pair (italicized). Participants only viewed one version of each sentence. The boundary was located immediately after the first lexeme/ adjective. Prior to fixating the second lexemes readers received either a correct preview or a partial preview (in parentheses).

1. Gary found an old tennis/ *filthy* ball (badk) in the corner of the garage.
2. The lights in the locker/ *empty* room (rosw) were annoyingly bright.
3. Fred wanted a picture of the mountain/ *Siberian* range (raoyz) for his collection.
4. George didn't expect the pillow/ *vicious* fight (fiqbk) in the attic to be so noisy./ *on the field to last so long.*
5. He wanted to buy/ *remove* a guitar/ *single* string (stucap) but he didn't have enough time.
6. The museum had a picture of a dinner/ *French* table (tahkm) from the late 18th century.
7. The antique coffee/ *wooden* table (tahkm) in the dining room was very expensive.
8. John pulled out a piece of carbon/ *yellow* paper (paqmo) from his desk drawer.
9. They installed a new garage/ *larger* door (dosu) next to the police station.
10. He didn't expect the traffic/ *orange* light (liykb) to come up so suddenly/ *to bother people so much.*
11. Mary was pleased to see health/ *fresh* food (fomt) in the new corner store.
12. The noise coming from the assembly/ *farthest* line (lioc) was surprisingly loud.
13. He wondered whether the exchange/ *original* rate (rafx) would change before the end of the day.

14. The visitors to the former prison/ *highly famous* camp (cavg) were very moved./ *were disappointed.*
15. For every beauty/ *active* queen (qucmo), morning exercises are a part of the daily routine.
16. Joan knew that one problem/ *naughty* child (chrkf) in the class could spoil it for everyone.
17. Mike needed to repair the kitchen/ *rusting* sink (sirl) before the week-end.
18. The experienced travel/ *former* agent (agmaf) liked to keep a diary of the places he visited.
19. The people in the concert/ *splendid* hall (hatf) relaxed and enjoyed the music.
20. The committee saw that the tennis/ *local* court (comvd) badly needed some work./ *badly needed a new judge.*
21. Linda thought that her new credit/ *cheery* card (caeh) was an improvement over the old one.
22. Johnny wondered who left the dollar/ *unpaid* bill (bifh) on the kitchen counter.
23. A flash of light appeared in the mirror/ *blurry* image (imnjw) that Ralph was staring at.
24. Nancy was fumbling with the rubber/ *sticky* band (baof) because it was so dark./ *because it had to go.*