ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Memory and Language 48 (2003) 615–634

Journal of Memory and Language

www.elsevier.com/locate/jml

The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds[☆]

Raymond Bertram* and Jukka Hyönä

Department of Psychology, University of Turku, FIN-20014 Turku, Finland

Received 25 June 2002; revision received 10 September 2002

Abstract

This study explored whether the length of a complex word modifies the role of morphological structure in lexical processing: Does morphological structure play a similar role in short complex words that typically elicit one eye fixation (e.g., eyelid) as it does in long complex words that typically elicit two or more eye fixations (e.g., watercourse)? Two eye movement experiments with short vs. long Finnish compound words in context were conducted to find an answer to this question. In Experiment 1, a first-constituent frequency manipulation revealed solid effects for long compounds in early and late processing measures, but no effects for short compounds. In contrast, in Experiment 2, a whole-word frequency manipulation elicited solid effects for short compounds in early and late processing measures, but mainly late effects for long compounds. A race model, incorporating a headstart for the decomposition route, in case whole-word information of complex words cannot be extracted in a single fixation can explain the pattern of results.

Keywords: Morphological processing; Eye movements; Word length; Finnish; Compound words

As many studies have shown, several factors can affect the way morphologically complex words are processed (Bertram, Laine, & Karvinen, 1999; Laudanna & Burani, 1995). Generally, very few of these factors are taken into account in psychomorphological research. This might be one reason why empirical evidence for the role of morphology in lexical processing is highly contradictory (see McQueen & Cutler, 1998, for an overview). Another reason for the contradictory results might be linked to one potential factor that has not been considered so far, namely word length. Naturally, word length as such affects processing speed (Just & Carpenter, 1980; Rayner & McConkie, 1976), but the question we would like to address here is whether the length of a complex word modifies the role of morphological structure in lexical processing. In other words, does morphological structure play a similar role in short morphologically complex words like EYE/LID as it does in long morphologically complex words like WATER/COURSE? From the eye movement literature (see Rayner, 1998; for a survey) it is known that the former type of words typically elicits one eye fixation, whereas the latter typically elicits two or more eye fixations. What may be derived from this is that the first constituent of a long compound might have a—what we will call—visual acuity benefit over the second constituent or over the whole word, whereas for short compounds, all

☆ This study was financially supported by the Finnish Cultural Foundation (Suomen Kultuurirahasto, grant to the first author) and the Academy of Finland (grant to the second author). We gratefully acknowledge John Henderson, Keith Rayner, and an anonymous reviewer for their valuable comments on an earlier version of this paper.

* Corresponding author. Fax: +358-2-3335060.
E-mail address: rayber@utu.fi (R. Bertram).

0749-596X/02/$ - see front matter © 2002 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0749-596X(02)00539-9

the sublexical and lexical information can be extracted at once. Consequently, the first constituent might simply participate in lexical access of long compounds, because it is in sharp focus during the first fixation, rather than for purely structural reasons. An earlier study on long (about 13 characters) Finnish compounds (Hyönä & Pollatsek, 1998) clarified the fact that the first constituent is clearly involved in lexical access from the very first fixation onwards. If this involvement is structural, we should find the same effects for short compounds. If, however, the effects are triggered by the visual acuity constraints of the eye, the role of the first constituent should be at least attenuated for short compounds. Like Hyönä and Pollatsek (1998) and Pollatsek, Hyönä, and Bertram (2000), we investigated the role of the first constituent and that of the whole word in compound processing by manipulating first-constituent and whole-word frequency, respectively. Many studies have demonstrated that frequency is one of the most robust factors driving processing speed. High-frequency words elicit shorter first fixation and gaze durations than low-frequency words (e.g., Henderson & Ferreira, 1990, 1993; Hyönä & Olson, 1995; Inhoff & Rayner, 1986; Just & Carpenter, 1980; Raney & Rayner, 1995; Rayner & Duffy, 1986). Therefore, studies on morphological processing have often employed frequency as a diagnostic tool. More precisely, manipulations of the frequency of the whole word and of morphological components have been used extensively to examine the role of morphology in complex word processing (e.g., Bertram, Hyönä, & Laine, 2000a; Bertram, Laine, Baayen, Schreuder, & Hyönä, 2000b; Bertram, Schreuder, & Baayen, 2000c; Bradley, 1980; Burani & Caramazza, 1987; Colé, Beauvillain, & Segui, 1989; Taft, 1979).

Research on compound words

In the following, we present a survey of previous research in which the role of constituents in processing compound words has been studied. The number of studies on this topic is relatively small compared to the number of studies that have dealt with inflectional and derivational morphology, even though in many languages compounding is the most productive word formation type. In a language like Finnish, compounds account for more than 50% of the word stock.[1] Perhaps the reason for the relatively small number of compound studies is that—as Libben (1998) puts it—of all polymorphemic word types compounds are most sensitive to semantic drift. Consequently, the few existing studies focus fairly often on the difference in processing transparent and partly or wholly opaque compounds (e.g., Monsell, 1984; Sandra, 1990; Zwitserlood, 1994; Jarema, Busson, Nikolova, Tsapkini, & Libben, 1999). These studies—all priming studies—almost unequivocally show that, at some level of representation, morphological constituents are involved in processing transparent compounds (see also Lima & Pollatsek, 1983). There seems to be some agreement that, on a nonsemantic level (orthographic or separate morphological level), even opaque compounds are represented and processed via their constituent morphemes (Jarema et al., 1999; Zwitserlood, 1994). Zwitserlood found that with a stimulus onset asynchrony (SOA) of 300 ms completely opaque Dutch compounds like KLOK/HUIS (which means 'core of an apple,' though taken literally it would be 'clock house') prime their first and second constituents equally as well as transparent (KERK/ORGEL 'church organ') or partly opaque (DRANK/ORGEL 'heavy drinker,' literally 'beverage organ') compounds. Jarema et al. showed for French opaque compounds that it works the other way around as well. With a SOA of 150 ms, they found that compounds' first and second constituents primed opaque compounds and transparent or partly opaque compounds equally well. In contrast to these studies, Zhou and Marslen-Wilson (2000) found evidence for a whole-word account for English using a cross-modal priming paradigm. More precisely, they found no priming between first and second constituents of compounds (thus BATH did not prime TUB), nor did compounds with a shared first constituent prime each other (thus HEADACHE did not prime HEADSCARF). However, the latter observation could be interpreted as inhibition between two second constituents that are competing for linkage to the same first constituent and is thus not evidence against separate access of the first or second constituent as such.

Using a semantic priming paradigm, both Zwitserlood (1994) and Sandra (1990) showed that transparent compounds are even related at a semantic level to their constituents. Zwitserlood found that transparent compounds like KERK/ORGEL 'church organ' primed words that were semantically related to the first constituent (PRIESTER 'priest') or second constituent (MUZIEK 'music'). Sandra did it the other way round and showed that words like KOE 'cow' and TAK 'branch' primed words like MELK/FLES 'milk-bottle' and FRUIT/BOOM 'fruit-tree' very well. For partly or wholly opaque compounds, the story is less clear. Whereas Zwitserlood found semantic priming effects for partly but not wholly opaque compounds with a SOA of 300 ms, Sandra did not find an effect for a mixture of partly and wholly opaque compounds with, in effect, a much longer SOA. The mixture of compound types in Sandra's study does not allow direct comparison of the results of both studies, but a closer inspection of his

[1] Our database of the second largest newspaper in the country (Turun Sanomat) comprises 1.48 million word types. Almost 800,000 of them (53.1%) are compounds.

compounds reveals that the majority are only partly opaque. Therefore, it is most likely that the longer SOAs in Sandra's study resulted in more time for nonrelated constituent meanings to decay or to be suppressed.

In studies with an aphasic population, the general tendency is also that morphological structure is of importance in processing compounds. Hittmair-Delazer, Andree, Semenza, and De Bleser (1994) found, in a picture-naming task, that German-speaking aphasics retained the compound structure in their output, even though answers were often incorrect. In addition, the incorrect answers often contained one constituent of the compound. Similar results were obtained in a study by Semenza, Luzzatti, and Carabelli (1997) with Italian aphasics, and in a study by Mäkisalo, Niemi, and Laine (1999) with a Finnish aphasic. Mäkisalo et al. also found evidence that the first constituent serves as the primary access code in processing compounds. More specifically, more existing words were created when the first constituent was retained and more neologisms when the second constituent was retained. Moreover, compound structure remained intact more often when the first constituent was named correctly than when the second constituent was named correctly. Libben (1993) reports a case study of an English-speaking patient, who was submitted to a paraphrase and a lexical decision task with 120 compound nouns. The former task yielded mostly transparent readings, even for opaque compounds (BLUEPRINT: 'a print that is blue'), but occasionally also blends of whole-word and constituent meanings or appropriate whole-word paraphrases. Interestingly, the latter mainly occurred when one of the constituents was low in frequency or phonologically obscure (the phonological form of the constituent was different from its phonological form as a free morpheme). The lexical decision task yielded further evidence that information about the whole-word form or meaning of compounds was maintained, even though paraphrase output was based mainly on the meanings of the constituents. The pattern of results led Libben to conclude that the patient's behavior "presents evidence in favor of a parallel access model of multimorphemic word recognition in which constituents and whole-word forms are simultaneously accessed."

This conclusion goes against the spirit of the seminal study of Taft and Forster (1976) with a normal population. In this study, four experiments with nonwords and one with real words support the view that compounds are decomposed before lexical access, and that the first syllable possibly coinciding with the first constituent serves as a lexical entry. Andrews (1986) partly replicated the findings of Taft and Forster (1976), but showed and argued convincingly that it is the first constituent rather than the first syllable that affects reaction times on compound words. In addition, she found evidence for a significant role of the second constituent in processing compounds, as did Lima and Pollatsek (1983). In the study of Lima and Pollatsek nonwords comprising two existing lexemes were more difficult to reject than nonwords that contained a nonexisting second lexeme. In Taft and Forster's (1976) study exactly the same comparison yielded no difference at all. Additional evidence for an important role of the second constituent comes from a study by Juhasz, Starr, Inhoff, and Placke (in press), who reported that compounds with high-frequency second constituents elicited faster lexical decision times, naming latencies and gaze durations than compounds with low-frequency second constituents. Manipulation of first-constituent frequency elicited only marginal tendencies, which led the authors to assign a more prominent role to the second constituent than to the first one. This is completely opposite to the findings of Taft and Forster's (1976) study, but it may not appear so surprising if one realizes that, generally, in English the second constituent is more closely related in meaning to the full compound than the first constituent. However, the weak frequency effects could also be partly due to the selection of low-frequency first constituents (e.g., beef, lamp, star, wash, and flash) that are more frequently in use in daily life than a written corpus would indicate.

Inhoff, Radach, and Heller (2000) presented German compounds with three constituents (e.g., DATA/SCHUTZ/EXPERT 'data protection expert') to native speakers in various ways: concatenated, with capitals for the first letter of each constituent, and with spaces between the constituents. It should be noted that the latter way of presenting compounds goes against standard German orthographic rules, but nevertheless, participants benefited significantly from the inclusion of spaces in a naming and in a normal reading task. What can be said then is that segmentation of a concatenated compound takes time. At any rate, all these studies more or less agree that initial access of compound words is morphologically motivated.

In contrast, on the basis of a series of lexical decision experiments in Dutch, Van Jaarsveld and Rattink (1988) found evidence for a whole-word access account of lexicalized compounds. A first constituent frequency effect was found for novel compounds, but for the lexicalized compounds only whole-word frequency effects were observed. Similarly, De Jong, Feldman, Schreuder, Pastizzo, and Baayen (2001) found no effect of constituent frequency in a visual lexical decision task but a significant whole-word frequency effect. In addition, they observed an effect of position-related morphological family for the first and second constituents. Derived and compound word formations starting with the same first constituent or ending with the same second constituent as target compounds are counted as position-related morphological family members. It should be noted that whole-word frequency was not orthogonally manipulated in their experiments, so the whole-word frequency effect they found is a correlational one. This effect might be a mere coincidence, since the whole-word frequency range of selected items was very limited (from 0 to 0.5). As they state themselves in the discussion, the effect of position-related morphological family might be a strategic one, since all the experiments consisted of compounds only. This state of affairs could lead to a situation in which 'participants made use of this kind of position-dependent probabilities to speed up their responses.' Be that as it may, their theoretical position is that compounds are accessed as whole units and that, at a later, more central level, compounds can benefit from having morphological relations with certain types of morphological family members.

In between these two camps is the position of Pollatsek and colleagues, who claim, on the basis of their results, that a morphological race is going on, in which efforts are made to access compounds simultaneously via a whole-word unit and in a decomposed format (Hyönä & Pollatsek, 1998; Pollatsek et al., 2000). In their studies, first constituent, second constituent, and whole-word frequency were manipulated in order to get a deeper insight into the processing of Finnish compound words. Importantly, in these studies, unlike in many other studies, compounds were included in sentences and eye movements were recorded, to obtain a detailed picture of the temporal characteristics of compound word processing. The first issue is important, since, naturally, we typically encounter words in sentences, which, among other things, allows us to partly process them before fixating on them (Rayner, 1998). Second, especially for long compounds, temporal characteristics are interesting, since the involvement of morphology can differ in early and late stages of processing. In other words, by employing the eye-tracking method, a clearer picture can be achieved about *when* exactly the first or second constituent or the whole-word form is involved in processing, if at all. This is something that cannot be done with, for instance, the visual lexical decision paradigm. Pollatsek and colleagues found the following pattern of results. For the first fixation duration, reflecting early stages of processing, there was a significant effect of first-constituent but not of second-constituent frequency. The effect of whole-word frequency came out significantly in the participant analysis with a slight tendency in the item analysis. For gaze duration, a measure that includes all fixations before exiting the word, there were solid effects of all frequency manipulations. The results are summarized in Fig. 1.

This pattern of results led to the conclusion that there is a clear involvement of the first constituent in early stages of processing compounds, but not of the second constituent and only mildly of the whole-word form. However, there seems to be a clear involvement of all of these elements in later stages of processing. On the basis

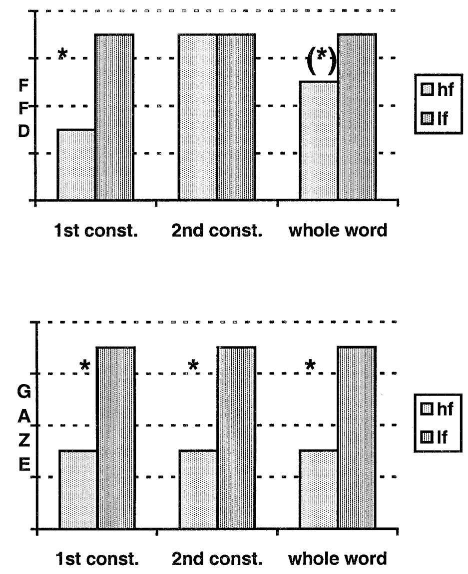

Fig. 1. Summary of the compound word results of Hyönä and Pollatsek (1998) and Pollatsek et al., 2000 with manipulations of the first-constituent frequency, second-constituent frequency, and whole-word frequency, respectively. Dependent measures include first fixation duration (FFD) and gaze duration (Gaze). The asterisks indicate significance at the 5% level; the asterisk in parentheses indicates a tendency toward significance.

of these results, Pollatsek and colleagues proposed a simple race model, in which both decomposed access and whole-word access of compounds is attempted. A recent eye movement study of Andrews, Miller, and Rayner (in press) showed that the findings for Finnish compound words generalize to English. They found frequency effects of the first and second constituents and the whole word in gaze duration and total fixation time, but only a hint of a whole-word frequency effect in the first fixation duration.

In sum, earlier studies leave us with the impression that the disagreement about the way compounds are represented and processed mirrors the disagreement that is present in the domains of inflection and derivation (see McQueen & Cutler, 1998, for a survey). More precisely, also in these domains decomposition accounts (e.g., Taft, 1994), whole-word accounts (Butterworth, 1983; Giraudo & Grainger, 2000) and race accounts are suggested (Baayen, Dijkstra, & Schreuder, 1997; Bertram et al., 1999; Schreuder & Baayen, 1995), alongside distributed connectionist approaches (e.g., Plaut & Gonnerman, 2000; Raveh & Rueckl, 2000), in which morphological structure is treated as an emerging property of the systematic relation between word forms and meanings. At any rate, most studies on compounds reveal that constituents play an important role in the course of processing. Some studies (Mäkisalo et al., 1999; Taft & Forster, 1976) indicate that the first constituent plays a more prominent role than the second constituent. In addition to the role of constituents, the Finnish studies of Pollatsek and colleagues, as well as the eye movement study of Andrews et al. (in press) show that the whole-word form is also involved in processing compounds be it that its significance seems to appear at later stages of processing. However, since compounds are constructed out of at least two free morphemes, they can be of considerable length. Because words of considerable length typically elicit two or more fixations, one may start to wonder whether the predominant involvement of constituents is purely structural in nature or whether word length as such drives the system towards a more pronounced use of morphological structure. In the following section we review some studies that indicate that the latter might be the case.

The interaction of word length and morphology

One study in which frequency effects might be confounded with word length is that of Colé et al. (1989). The authors find effects of base frequency for their suffixed words but not for their prefixed words. Accordingly, they claim that suffixed but not prefixed words are accessed via their root. However, upon closer inspection, it can be seen that the words in the four prefixed conditions have a length of around 8 letters, whereas the words in the four suffixed conditions average around 9.5 letters. Even though this might not seem such a big difference, the saccade length in French is typically 8 letters (Levy-Schoen & O'Regan, 1979). In other words, whereas the prefixed conditions can fairly often be dealt with in a single eye fixation, the suffixed ones would more frequently need two eye fixations. This state of affairs could have enhanced the decomposition route in case of suffixed words, and the direct route in case of prefixed words. Thus, instead of the processing system making a principled distinction between prefixed and suffixed words, a simple physical distinction might be at stake here. This suspicion is aroused even more by a subsequent eye movement study of Beauvillain (1996) on French prefixed and suffixed words. This time, she observed a base frequency effect of both suffixed and prefixed words, when words with a length of 10 characters or more were selected. Interestingly, the effect for suffixed words, with the stem in the beginning of the word, is found on the first fixation. In contrast, the effect for prefixed words, with the stem further in the word, is observed on the second fixation and in the gaze duration. In addition, whole-word form effects only emerged in the second fixation duration and in the gaze duration. This pattern of results suggests that the role of morphology in longer words may be determined at least partly by visual constraints of the eye.

In a similar vein, Niswander, Pollatsek, and Rayner (2000), in their first experiment, found a solid base frequency effect for long derivations (>9.5) in the first fixation duration and a solid surface frequency effect in both the second fixation duration and the gaze duration. For the much shorter inflections (<7.0), only a whole-word frequency effect was observed. This effect appeared in the first fixation duration. Again, this study is in line with the idea that word length may significantly modify the role of morphology in lexical processing.

However, definite claims about the possible interaction of morphology and word length can be made neither on the basis of a comparison between the Colé et al. (1989) and the Beauvillain (1996) study, nor on the basis of a comparison between Experiments 1 and 2 of the Niswander et al. (2000) study. First, the complex words used in Experiments 1 and 2 of the latter study are of a different nature. The word types used in Experiment 1 are derivations, while in Experiment 2, a variety of inflectional types was employed. Second, the frequency values of Experiments 1 and 2 were not matched. The same holds for the short and long prefixed words in Colé et al. (1989) and Beauvillain (1996).[2] To compare the exact interaction between morphology and word length, the same types of long and short complex words should be investigated and frequency values for long and short conditions should be kept constant. An additional concern about the French studies is that they both employed a nonreading paradigm (lexical decision and semantic judgment task), which leaves unanswered the question of how short and long complex words are processed under more natural circumstances when they are presented in context.

The present compound study

In the current study, we investigated the role of morphology in short and in long compounds in context. In Experiment 1, we manipulated the frequency of the first constituent in an identical way for long compounds (⩾12) like JOUKKUE/HENKI 'team spirit' and short compounds (⩽8) like JÄÄ/RATA 'ice ring,' while keeping (among other factors) average second-constituent frequency and average whole-word frequency constant for all four conditions. In Experiment 2, we manipulated the frequency of the whole-word form for long and short compounds in a similar vein, while matching for average second constituent and average whole-word frequency across all four conditions. Several studies show that frequency is one of the most powerful factors to exert differences in word processing times (Jescheniak &

[2] This is not a weakness of these studies. They were simply not set up to directly test the role of word length on morphological processing.

Levelt, 1994; Rubenstein & Pollack, 1963; Whaley, 1978). In the studies examining the processing of morphologically complex words, frequency is often used as a diagnostic tool. More specifically, the frequency of the constituent morphemes or that of the whole word is manipulated in order to assess whether morphological substrings and/or whole-word forms are employed in the course of processing. Under some circumstances, both types of manipulation exert an effect (e.g., Bertram et al., 2000a,b,c; Burani & Caramazza, 1987; Taft, 1979), whereas under other circumstances only one of them does (e.g., Bertram et al., 2000a,b,c; Bradley, 1980; Colé et al., 1989; Vannest & Boland, 1999). As in previous studies, we assume that if low first-constituent frequency compounds take longer to read than high-frequency ones, this would imply that the first constituent is effectively employed in accessing the whole compound. On the other hand, if the frequency manipulation of the first constituent does not exert an effect, access is not assumed to take place via the first constituent. Similarly, if low whole-word frequency compounds take longer to read than high-frequency ones, it would imply that the whole-word form is effectively employed in accessing the compound. If the whole-word frequency manipulation does not exert an effect, access via the whole-word form is assumed not to take place.

In both experiments, we paired high- and low-frequency items and presented them in a sentence frame that was identical through the word following the target word. This allowed us to reliably assess the effects of frequency and word length not only on the target word, but also on the word following the target word, so that possible spillover effects could be taken into account when interpreting the data.

Finally, in both experiments, we used readers' eye movements as the measure of on-line processing. Eye-tracking has now become the most preferred method of studying on-line visual language processing (see Rayner, 1998). One of the major advantages of this method is that similarly to daily reading it allows readers to freely inspect the text in the way they wish. In addition, the method allows a temporal insight into word processing, which is particularly attractive for the type of compound study we conducted.

Experiment 1

Method

Participants. Thirty students of the University of Turku participated in the experiment. All were native speakers of Finnish, and had normal or corrected-to-normal vision.

Apparatus. Eye movements were monitored by the EYELINK eyetracker manufactured by SR Research (Canada). The eyetracker is an infrared video-based tracking system combined with hyperacuity image processing. There are two cameras mounted on a headband (one for each eye) including two infrared LEDs for illuminating each eye. The headband weighs 450 g in total. The cameras sample pupil location and pupil size at the rate of 250 Hz. Registration is monocular and is performed for the selected eye by placing the camera and the two-infrared light sources at a distance of 4–6 cm from the eye. The spatial accuracy is better than 0.5°. The spatial resolution (i.e., the differential accuracy) of the system is 15 min of arc. Head position with respect to the computer screen is tracked with the help of a head-tracking camera mounted on the center of the headband at the level of the forehead. Four LEDs are attached to the corners of the computer screen, and are viewed by the head-tracking camera when the subject sits directly facing the screen. Possible head motion is detected as movements of the four LEDs and is compensated for online from the eye position records. The system allows free head motion within a 100 cm cube.

Materials. Forty short and 40 long two-noun compounds were selected from an unpublished computerized newspaper corpus of 22.7 million word forms with the help of the WordMill database program of Laine and Virtanen (1999). Twenty long compounds had a high-frequency first constituent with a mean of 472 occurrences per million (range 72–1052), whereas the other twenty had a (relatively) low-frequency first constituent with a mean of 23 per million (range 1–117; all frequency counts reported are scaled to 1 million). Twenty short compounds had a high-frequency first constituent with a mean of 468 occurrences per million (range 72–1186), whereas the other 20 had a (relatively) low-frequent first constituent with a mean of 25 per million (range 1–101).[3] To find out whether this manipulation was strong enough to exert an effect, and also to find out whether it was similar for long and short compounds, we conducted a familiarity rating task with all 80 target words. Ten participants who did not participate in Experiment 1 proper rated how familiar each word was for them on a scale from 1 to 5 (1 = unfamiliar; 5 = very familiar). The long and short compounds with a high-frequency first constituent were rated with a mean of 3.0, those with a low-frequency first constituent were rated with a

[3] For the long compounds the low-frequency item with 117 occurrences per million was the only item that overlapped with the high-frequency condition. The item itself was paired with a high-frequency item that had a first constituent occurring 737 times per million. For the short compounds, two low-frequency items with 93 and 117 occurrences per million, respectively, were the only items that overlapped with the high-frequency condition. The items themselves were paired with high-frequency items that had a first-constituent frequency of 487 and 180 per million, respectively.

mean of 2.5 and 2.6, respectively. This yielded a solid frequency effect ($F1(1,9) = 18.3$, $p < .001$; $F2(1,38) = 7.2$, $p < .02$), but no effect of word length or Frequency × Word Length interaction (all Fs < 1). In other words, the frequency manipulation had been successful and was similar for long and short compounds.

The high- and low-frequency conditions were matched for surface frequency,[4] second constituent frequency, bigram frequency, word length in letters, and length of the first constituent. The short compounds were closely matched with the long compounds in all aspects except (naturally) constituent and word length. The lexical–statistical properties of all conditions are listed in Table 1.

The target words were embedded in sentences with each target word appearing in a separate sentence. Each of the compound words starting with a high-frequency first constituent was paired with a compound word with a low-frequency first constituent, and a sentence frame was constructed that was identical up to the word following the target word; the rest of the sentence was different. To match for semantic plausibility, a rating study was conducted, in which both versions of the sentence pairs were listed underneath each other, and seven participants who did not participate in the experiment proper rated the naturalness of the sentences of Experiments 1 and 2 by using one of three possible alternatives: sentence 1 sounds more plausible, sentence 2 sounds more plausible, or sentences 1 and 2 sound equally plausible. If at least 3 out of 7 rated one sentence to be more plausible than the other, a new sentence frame was constructed. This only happened on a few occasions (4 out of 40 in Experiment 1, and 5 out of 40 in Experiment 2). Another set of 13 individuals, who did not participate in the experiments proper or in any of the rating tasks, was given the target sentence frames of Experiments 1 and 2 up to the target word and asked to write down a word that could complete the sentence fragment. None of the target words was generated as completions of the sentence fragments and, in general, very few compound words were generated. In sum, the final sets of target words were equally plausible, but not predictable from the preceding sentence context. An example of a sentence pair from Experiment 1 is shown below (the target word is shown in bold).

[4] Inflectional variants of a compound word (e.g., JÄÄ/RATA/A 'ice ring + partitive' or JÄÄ/RATA/A/KIN 'ice ring + partitive + clitic denoting also' of the nominative singular JÄÄ/RATA) are also included in the counts of surface frequencies as reported in Experiments 1 and 2. Thus, in principle, one could say that the reported surface frequencies are actually lemma frequencies. Note, however, that the items were matched (Experiment 1) or manipulated (Experiment 2) on surface frequency in the strict sense (thus only on the nominative singular case) as well.

Low-frequency first constituent:

*Miehen omituinen **lannevyö** sai kadunkulkijoiden katseet kääntymään.*

'The man's peculiar **loin belt** made the pedestrians turn their heads.'

High-frequency first constituent:

*Miehen omituinen **työpuku** sai uteliaat työtoverit ihmettelemään sen alkuperää.*

'The man's peculiar **working outfit** made his curious colleagues wonder about its origin.'

The target sentences were presented in Courier one at a time, starting from the center left position on the computer screen. The sentences took up a maximum of three lines of text, the critical word never appearing as the initial or final word of a text line. With a viewing distance of about 65 cm, one character space subtended approximately 0.5° of visual angle. The 80 target sentences were mixed with 64 filler sentences. The sentences were presented in two blocks, so that paired sentences never appeared in the same block. The order of the blocks was counterbalanced across participants, and within a block the order of sentences was randomized.

Procedure. Prior to the experiment, the eyetracker was calibrated using a nine-point calibration grid that extended over the entire computer screen. Prior to each sentence, the calibration was checked by presenting a fixation point in the center left position of the screen; if needed, calibration was automatically corrected, after which a sentence was presented to the right of the fixation point.

Subjects were instructed to read the sentences for comprehension at their own pace. They were further told that periodically they would be asked to paraphrase the last sentence they had read to make sure that they were attending to what they read. It was emphasized that the task was to comprehend, not to memorize the sentences. Subjects were asked to paraphrase a sentence approximately after every seven sentences. The experimental session lasted a maximum of 50 min.

Results

Separate analyses were conducted for the target word and for the word following the target word ($N + 1$). Individual fixation times shorter than 50 ms were excluded before analyses. For the target word, several dependent measures were analyzed of which five are reported and discussed extensively: gaze duration, initial landing position, first fixation duration, saccade length from the first to the second fixation location (hereafter referred to as saccade length), and the probability of

Table 1
Lexical–statistical properties of the four conditions in Experiment 1

	Long compounds with high-frequency first constituent	Long compounds with low-frequency first constituent	Short compounds with high-frequency first constituent	Short compounds with low-frequency first constituent
Mean first-constituent frequency[a]	472	23	468	25
Mean second-constituent frequency[a]	242	237	250	238
Mean surface frequency[a]	2.7	2.9	3.1	2.5
Mean familiarity rating[b]	3.0	2.5	3.0	2.6
Mean bigram frequency[c]	7.6	7.3	6.7	5.5
Mean word length[d]	12.7	12.9	7.6	7.8
Mean first constituent length[d]	6.7	7.1	3.8	4.1

[a] All values scaled to 1 million.
[b] Rating scale from 1 to 5.
[c] Scaled to 1000.
[d] Word length in characters.

making a third fixation.[5] These measures reflect different stages of processing: preprocessing, early processing, and late processing. Considered jointly, they give a detailed temporal insight into lexical processing. By analyzing initial landing position, one could discern whether preprocessing of the target word, that is, processing of the compounds before actually fixating it, has taken place. Differences in initial landing position would imply that some aspect of the target word (first-constituent frequency and/or word length) triggers a longer or shorter saccade from a previous word into the target word. Two measures that reflect early stages of target word processing are first fixation duration and intraword saccade length. First fixation duration is an early processing measure as it indexes the initial encounter with the word. Saccade length reflects early processing, since the 'decision' to make a second fixation and where to locate it takes place during the first fixation. Then, the probability of a third fixation reflects processing efforts after the initial stages. Finally, gaze duration is the summed duration of all fixations on the target word before fixating away from it. Thus, it is a composite score of individual fixation durations and within-word refixation probability, and it provides a measure of the total time needed to process a word during the first pass reading.

For all target word measures, analyses of variance (ANOVA) were conducted with both participants and items as the random variable. For the participant analyses, frequency and word length were treated as within-participant factors, while for the item analyses, frequency was treated as a within-item factor and word length as a between-item factor. Interactions were further examined by two-tailed t tests. The averages of the five reported measures (plus second fixation probability) for the long and short compounds are presented in Table 2.

In order to assess possible spillover effects in processing, first fixation duration and gaze duration were analyzed for the $N + 1$ word. However, there were no reliable spillover effects in either measure, indicating that all possible morphological effects were revealed on the target words themselves. The same applied for Experiment 2.

Target word

Gaze duration. Not surprisingly, a main effect of word length was found ($F1(1, 29) = 255.4$, $p < .001$; $F2(1, 38) = 118.4$, $p < .001$), indicating that longer compounds were gazed upon longer than short compounds. The main effect of frequency was also significant, indicating that compounds with a high-frequency first constituent needed less gaze time than compounds with a low-frequency first constituent ($F1(1, 29) = 32.9$, $p < .001$; $F2(1, 38) = 10.0$, $p < .005$). In addition, the Word Length × Frequency interaction was significant in both analyses ($F1(1, 29) = 15.2$, $p = .001$; $F2(1, 38) = 5.4$, $p < .03$). Consequently, we performed a set of follow-up t tests to compare the high- and low-frequency conditions separately for long and short compounds. By doing so, a clear frequency effect for long compounds was observed. Long compounds with a high-frequency first

[5] Other measures that were also analyzed include skipping rate, probability of two fixations, probability of four fixations, second fixation duration and second fixation location. Since these measures did not provide any major additional insight into the time course of processing Finnish compounds, we do not report them here, although sometimes they feature in a footnote. Another measure that is not reported is the total time that is spent fixating the compound word (including regressions). This measure was left out, as it would not give additional insight into whether or not our foveal visual acuity benefit hypothesis holds.

Table 2
Gaze duration (in ms), initial landing position (in character spaces), first fixation duration (in ms), saccade length (in character spaces), and probability of a second and third fixation for the long and short compounds with high- vs. low-frequency first constituents of Experiment 1 (with all standard deviations in parentheses)

Eye movement measure	Long compounds		Short compounds	
	High	Low	High	Low
	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
Gaze duration	472 (61)	542 (83)	336 (44)	347 (56)
Initial landing position	4.23 (.90)	4.21 (.98)	3.43 (.63)	3.49 (.61)
First fixation duration	212 (21)	228 (22)	246 (30)	243 (32)
Saccade length	4.16 (.92)	3.60 (.93)	–	–
Probability of a second fixation	.833 (.09)	.848 (.07)	.369 (.16)	.403 (.14)
Probability of a third fixation	.330 (.14)	.472 (.14)	.057 (.05)	.075 (.07)

constituent elicited shorter gaze durations than those with a low-frequency first constituent ($t1(29) = 5.3$, $p < .001$; $t2(39) = 3.0$, $p < .005$). However, the effect for short compounds failed to reach significance in the item analysis ($t1(29) = 2.6$, $p = .02$; $t2(38) < 1$). These results indicate that the role of the first constituent is more pronounced for long than for short compounds.

In subsequent analyses, we examined in more detail the time-course of the constituent frequency effect. We first report the measure indexing preprocessing, then the measures indexing early processing stages followed by the measure reflecting a relatively late stage.

Initial landing position. This measure indexes the target word processing prior to fixating it. The initial landing position was found to be located further in the word for long compounds ($F1(1,29) = 100.1$, $p < .001$; $F2(1,38) = 9.9$, $p = .003$). As found in earlier studies (e.g., O'Regan, 1992; Rayner, 1979), this result shows that low-level features of the upcoming word like word length are already processed parafoveally and used in the decision on where to locate the next eye fixation. However, no main effect of frequency or a Word Length × Frequency interaction was observed (all Fs < 1), indicating that first-constituent frequency did not affect parafoveal processing of short and/or long Finnish compounds. Next, we proceed to the analyses of measures that reflect early processing stages when the target word is fixated.

First fixation duration. For the first fixation duration, there was a main effect of word length ($F1(1,29) = 29.5$, $p < .001$; $F2(1,38) = 11.6$, $p < .005$), but this time longer compounds were initially fixated for a shorter time than short compounds. The frequency effect was significant in the participant analysis ($F1(1,29) = 4.1$, $p = .05$), but not in the item analysis ($F2(1,38) = 2.3$, $p = .14$). The Word Length × Frequency interaction was significant in both analyses ($F1(1,29) = 11.1$, $p = .002$; $F2(1,38) = 4.2$, $p < .05$). Subsequent t tests showed that there was a clear frequency effect for long compounds, high-frequency long compounds eliciting shorter first fixations than low-frequency compounds ($t1(29) = 5.5$, $p < .001$; $t2(38) = 2.5$, $p < .02$). However, this effect did not show up for the short compounds ($t1,2 < 1$). It is still possible that the differential frequency effect in the two types of compound might be a consequence of the differential nature of first fixations. Whereas, for long compounds, the first fixation is mostly followed by one or more fixations on the same word (in 84% to be precise), for short compounds this is not typically the case (here only in 39% of cases). Consequently, most of the first fixations on short compounds might be different in nature than most of the first fixations on long compounds. In order to check whether the type of first fixation may account for the null result for short compounds, we ran two follow-up analyses. For the first analysis, we selected those 39% of cases in which the short compounds had elicited at least two fixations. The average first fixation duration for the high-frequency condition in these instances was 220 ms, while for the low-frequency condition, it was 222 ms. The t tests showed that the conditions did not differ significantly from each other ($t1,2 < 1$). Second, we analyzed the 61% of instances with exactly one fixation. The average first fixation duration here was 263 ms for the high-frequency condition and 264 ms for the low-frequency condition. The t tests showed again that the conditions did not differ significantly from each other ($t1,2 < 1$).[6]

It is also possible that the lack of a first-constituent frequency effect in the first fixation duration of short compounds arose from the fact that the readers did not fixate on the first constituent in approximately 40% of the trials (in other words, in about 40% of the trials the landing position was on the second constituent). However, separate analyses including only trials in which the

[6] The roughly 40 ms increase in both conditions shows that in some sense the nature of the first fixation is different in single-fixation cases in comparison to multiple-fixation cases. However, this difference did not affect the first-constituent frequency effect.

landing position was on the first constituent, revealed no significant difference between the high frequency (238 ms) and low-frequency (240 ms) condition ($t1, 2 < 1$). Thus, we conclude that manipulation of first-constituent frequency did not affect the first fixation duration on short compounds, no matter how you look at it.

Saccade length. The second measure of early processing is that of intraword saccade length, which, in this case, is the saccadic amplitude between the first and second fixation locations. The problem with these analyses is that there is a considerable difference in the statistical power for the long compared to the short compounds. Whereas every participant fixated most of the long compounds at least twice, several participants fixated most of the short compounds only once. In addition, whereas each item in the long conditions elicited a second fixation for more than 60% of the time, practically all items in the short conditions elicited a second fixation for less than 50% of the time, many of them even less than 25%. This means that many of the item and participant means in the short conditions are based on only a few observations. Therefore, we decided to run the analyses for long compounds only. The separate *t* tests showed that the frequency effect for long compounds was a solid one ($t1(29) = 3.6$, $p = .001$; $t2(38) = 2.3$, $p < .03$).[7] The length of the intraword saccade was shorter when the first constituent was infrequent. Next, we report the analyses of the measure of a later processing stage.

Probability of third fixation. For the probability of a third fixation, there was again a significant main effect of word length ($F1(1, 29) = 97.0$, $p < .001$; $F2(1, 38) = 179.2$, $p < .001$; long words were more likely to be fixated three times than short ones. The main effect of frequency also proved significant ($F1(1, 29) = 33.2$, $p < .001$; $F2(1, 38) = 11.1$, $p < .005$), high-frequency words eliciting fewer triple fixations than low-frequency words. In addition, the Frequency × Word Length interaction proved significant ($F1(1, 29) = 15.1$, $p = .001$; $F2(1, 38) = 7.4$, $p = .01$). Subsequent analyses showed that, for long compounds, the probability of making a third fixation is higher for the low- than for the high-frequency condition ($t1(29) = 5.7$, $p < .001$; $t2(38) = 3.2$, $p < .005$). For short compounds the difference was not significant ($t1(29) = 1.4$, $p = .19$; $t2(38) < 1$).

Discussion

Frequency effects. The measure of preprocessing, the initial landing position, showed that first-constituent frequency did not exert an effect prior to fixating the target word. It is quite clear, however, that for the target word itself the pattern of results is different for long and short compounds. Whereas for long compounds, the first-constituent frequency manipulation elicited a clear frequency effect in first fixation duration, for short compounds an identical manipulation exerted no effect. On the basis of these results, we conclude that the first constituent is initially involved in processing long but not short compounds. The effect in first fixation duration for long compounds basically reflects a foveal effect: the high-frequency first constituent under foveal inspection is activated faster than the low-frequency first constituent.

The saccade length effect for long compounds is assumed to reflect an intraword preprocessing effect. In other words, there are more processing resources available to analyze the latter nonfoveal part of a compound word with a high-frequency first constituent. In effect, this may lead to a larger attentional span (by which we refer to the area from which useful letter information can be extracted) for long compounds with a high-frequency first constituent than for long compounds with a low-frequency first constituent, which in turn triggers longer within-word saccades in the former type (see, e.g., Henderson & Ferreira, 1990; Hyönä, 1995; Hyönä & Pollatsek, 2000). We discuss this issue in greater detail in the Discussion of Experiment 2.

In addition to these early effects of first-constituent frequency for long compounds, we also found an effect in the measure reflecting later stages of processing. The need for a third fixation was notably less when the first constituent of a long compound was of high frequency than when it was of low frequency. In other words, the frequency effect seems to linger on for long compounds after the first fixation.

Finally, the clear frequency effect in gaze duration for long compounds in comparison to the mild, if any, effect for short compounds indicates that, overall, the role of the first constituent is more pronounced for long than for short compounds.

The pattern of results for long compounds is compatible with models that insist on a primary role for the first constituent in lexical access (e.g., Mäkisalo et al., 1999; Taft, 1994) and those that insist on decomposed access in general (e.g., Zwitserlood, 1994), but against whole-word accounts (e.g., Giraudo & Grainger, 2000; Zhou & Marslen-Wilson, 2000). However, the pattern of results for short compounds seems incompatible with any model that advocates decomposition in the early stages of processing complex words.

Word length effects. It might not come as a surprise that long compounds were gazed upon longer than short compounds or that they elicited more second and third fixations. Nevertheless, we would like to point out that our premise that, in most cases, compounds of 8

[7] Another measure reflecting early processing, namely probability of second fixation, yielded neither a significant frequency effect ($F1(1, 29) = 3.1$, $p = .09$; $F2(1, 38) = 1.4$, $p = .24$) nor a significant Frequency × Word length interaction ($F1, 2 < 1$). Only the word length effect was highly significant ($F1(1, 29) = 179.2$, $p < .001$; $F2(1, 38) = 213.6$, $p < .001$).

characters or less can be dealt with in one fixation, while compounds of 12 or more characters typically elicit two or more fixations, was warranted. We conclude that the visual acuity of the end part of long compound words dropped off so dramatically on the first fixation that almost always a second fixation was needed to bring the end of the word into foveal vision for detailed analysis. For short compounds this was not the case. Nevertheless, even here a second fixation was sometimes needed, most often probably because whole-word access was slow. We suspect that, in these instances the motor program for a refixation was already in a nonlabile phase and could no longer be cancelled (cf. E-Z Reader 3 of Reichle, Pollatsek, Fisher, & Rayner, 1998). However, it should be noted that even when short compounds elicited multiple fixations, most often the whole word fell within the foveal span. The average initial landing position of the multiple fixation cases of both high-frequency and low-frequency short compounds was 3.1. Given that the average length of high-frequency short compounds was 7.6 and that of low-frequency ones 7.8, on average 4.5 and 4.7 characters, respectively, were left to be analyzed. The average initial landing position of the multiple fixation cases of both high-frequency and low-frequency long compounds was 4.2. Given that the average length of high-frequency long compounds was 12.7 and of low-frequency ones 12.9, on average, 8.5 and 8.7 characters, respectively, were left to be analyzed, part of which fell outside the foveal span. In short, there were twice as many letters to be analyzed on the right side of the fixation for multiply fixated long compounds than there were for multiply fixated short compounds.[8]

In short, the results of Experiment 1 lend support to our visual acuity hypothesis according to which (1) readers start to analyze the first constituent of long compounds simply because they do not have enough letter information on the latter part of the word and (2) readers start to analyze the whole-word string of short compounds immediately because all letters are within foveal vision. To find more direct support for the second part of this hypothesis, we examined the role of whole-word frequency in Experiment 2.

Experiment 2

Experiment 1 showed that for short compounds the role of the first constituent, if any, is late and not very pronounced. Consequently, one would expect that, for this type of compound, the role of the whole-word form would be of greater importance, playing an early and more influential role in word processing. In contrast, for long compounds one would expect that the role of the whole-word form would manifest itself later due to the nonoptimal visual acuity of the latter part of the word. The goal of Experiment 2 is to investigate whether these predictions are borne out by manipulating the whole-word frequency of long and short compounds, while matching for other relevant factors.

Method

Participants. Thirty students at the University of Turku participated in the experiment. None of them took part in Experiment 1 or in any of the rating tasks. All were native speakers of Finnish and had normal or corrected-to-normal vision.

Apparatus. The apparatus was identical to that used in Experiment 1.

Materials. Forty short and 40 long compounds were selected from our unpublished computerized newspaper corpus. Twenty long compounds were of (relatively) high frequency (23 occurrences per million, range 7–54), whereas the other twenty were of low frequency (2.4, range 0.6–6.7). Twenty short compounds were of (relatively) high frequency (22 occurrences per million, range 7–69), whereas the other twenty were of low frequency (2.3, range 0.5–5.7). Again, in order to find out whether these manipulations were strong enough to exert an effect and whether they were similar for long and short compounds, a familiarity rating task with all 80 target words was conducted. The same 10 raters as in Experiment 1, who did not participate in Experiment 1 or 2 proper, rated how familiar each word was for them on a scale from 1 to 5. The long and short compounds with a high-frequency whole-word form were rated with a mean of 3.6 and 3.9, respectively. Those with a low-frequency whole-word form were rated with a mean of 2.5. This yielded a solid frequency effect ($F1(1,9) = 296.9$, $p < .001$; $F2(1,38) = 100.0$, $p < .001$), but no effect of word length or Frequency × Word Length interaction (word length: both $F\text{s} < 1$; Word Length by Frequency: $F1(1,9) = 2.9$, $p > .1$; $F2(1,38) = 1.1$, $p > .3$). In other words, the frequency manipulation was successful and was similar for long and short compounds.

The high- and low-frequency conditions were matched on first-constituent frequency, second-constituent frequency, bigram frequency, word length in letters, and length of the first constituent. The short compounds were matched with the long compounds on all factors but constituent and word length. The lexical–statistical properties of all conditions are listed in Table 3.

Procedure. The procedure was identical to that of Experiment 1.

[8] On a few occasions, the initial landing position was in the very beginning of the compound word, far from the optimal viewing position (OVP). Accordingly, an immediate refixation was planned closer to this OVP (cf. O'Regan, 1992; Rayner, Sereno, & Raney, 1996).

Table 3
Lexical–statistical properties of the four conditions in Experiment 2

	Long compounds with high-frequency whole-word form	Long compounds with low-frequency whole-word form	Short compounds with high-frequency whole-word form	Short compounds with low-frequency whole-word form
Mean first-constituent frequency[a]	268	236	287	265
Mean second-constituent frequency[a]	280	270	284	314
Mean surface frequency[a]	25	2.3	22	2.1
Mean familiarity rating[b]	3.6	2.5	3.9	2.5
Mean bigram frequency[c]	7.9	7.0	5.9	6.1
Mean word length[d]	12.8	12.9	7.5	7.6
Mean first-constituent length[d]	6.7	7.0	3.7	3.8

[a] All values scaled to 1 million.
[b] Rating scale from 1 to 5.
[c] Scaled to 1000.
[d] Word length in characters.

Results

Similar analyses as in Experiment 1 were run for the target word. The means of the relevant measures for the target word can be found in Table 4.

Target word

Gaze duration. There was a main effect of word length ($F1(1,29) = 140.4$, $p < .001$; $F2(1,38) = 116.4$, $p < .001$), indicating that long compounds were gazed upon longer than short compounds. The main effect of frequency was significant ($F1(1,29) = 42.4$, $p < .001$; $F2(1,38) = 27.9$, $p < .001$), with high-frequency compounds being gazed upon for less time than low-frequency compounds. The Word Length × Frequency interaction did not reach significance in the item analyses ($F1(1,29) = 6.3$, $p < .02$; $F2(1,38) = 1.3$, $p = .27$). Apparently, the whole-word form plays an important role in processing both long and short compounds.

In subsequent analyses, we examined in more detail the time-course of the constituent frequency effect. We first report the measure indexing preprocessing, then the measures reflecting early stages of processing, followed by a measure reflecting a relatively late stage of processing.

Initial landing position. This measure indexes the target word processing prior to fixating it. As found in earlier studies and in Experiment 1, the initial fixation landed further into long than short words ($F1(1,29) = 57.2$, $p < .001$; $F2(1,38) = 13.0$, $p = .001$). There was no main effect of frequency ($F1(1,29) = 2.4$, $p = .13$; $F2(1,38) = 2.2$, $p = .15$). The Frequency × Word Length interaction remained nonsignificant in the item analysis ($F1(1,29) = 4.4$, $p = .05$; $F2(1,38) = 1.7$, $p = .21$), indicating once again that the effect of frequency prior to fixating the target word is at best minimal. We next report the analyses of measures that reflect early processing stages of the target word.

First fixation duration. There was a main effect of word length ($F1(1,29) = 24.7$, $p < .001$; $F2(1,38) = 9.5$, $p < .005$), indicating that, as in Experiment 1, long compounds were initially fixated for a shorter time than short compounds. The main effect of frequency was significant ($F1(1,29) = 11.5$, $p < .005$; $F2(1,38) = 6.7$, $p = .01$), with high-frequency compounds eliciting shorter first fixations than low-frequency compounds. The Word Length × Frequency interaction did not reach significance ($F1(1,29) = 2.0$, $p = .17$; $F2(1,38) = 1.3$, $p = .26$). This pattern of results leads to the conclusion that at first fixation, the whole-word form is involved in processing both short and long compounds.

Saccade length. Again, due to the difference in statistical power between long and short compound conditions, we only performed analyses for long compounds by means of t tests. In contrast to Experiment 1, no frequency effect for long compounds was found ($t1 < 1$; $t2 = 1.4$, $p = .17$).[9]

The final set of analyses was computed for a measure indexing a relatively late stage of processing.

Probability of third fixation. The probability of making a third fixation on the target word yielded a main effect of word length ($F1(1,29) = 84.7$, $p < .001$; $F2(1,38) = 88.8$, $p < .001$) and frequency ($F1(1,29) = 28.0$, $p < .001$; $F2(1,38) = 17.6$, $p < .001$). High-frequency compounds elicited fewer third fixations than

[9] Another measure reflecting early processing, namely, probability of second fixation, yielded a significant frequency effect ($F1(1,29) = 6.9$, $p = .01$; $F2(1,38) = 5.7$, $p = .02$), but no significant Frequency × Word Length interaction ($F1$, $2 < 1$). The word length effect was highly significant as well ($F1(1,29) = 120.8$, $p < .001$; $F2(1,38) = 137.2$, $p < .001$).

Table 4
Gaze duration (in ms), initial landing position (in character spaces), first fixation duration (in ms), saccade length (in character spaces), and probability of a second and third fixation for the long and short compounds with high- vs. low-frequency first constituents of Experiment 2 (with all standard deviations in parentheses)

Eye movement measure	Long compounds		Short compounds	
	High	Low	High	Low
	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
Gaze duration	400 (56)	479 (61)	300 (32)	352 (53)
Initial landing position	4.23 (1.05)	4.25 (.93)	3.23 (.63)	3.48 (.61)
First fixation duration	183 (17)	187 (19)	198 (22)	208 (23)
Saccade length	4.57 (.79)	4.30 (.97)	–	–
Probability of a second fixation	.815 (.08)	.847 (.08)	.456 (.16)	.525 (.10)
Probability of a third fixation	.299 (.14)	.429 (.16)	.072 (.07)	.147 (.07)

low-frequency ones; moreover, long compounds elicited more fixations than short ones. There was no clear evidence for a Word Length × Frequency interaction ($F1(1,29) = 3.1$, $p = .09$; $F2(1,38) = 1.2$, $p = .27$).

Long compounds: Pooled analyses of Experiments 1 and 2

As noted in the introduction to Experiment 2, we expected the role of the whole-word form of long compounds to manifest itself later than in the first fixation, because of the nonoptimal visual acuity of the latter part of the word. However, when analyzing long and short compounds together we found a main effect of frequency without any evidence of an interaction between frequency and word length. Thus, it seems that, contrary to our prediction, whole-word frequency is involved in the early processing of long compounds. What can be observed, however, is that the effect size for long compounds in Experiment 1 is larger than in Experiment 2: 16 ms vs. 4 ms.[10] To explore whether the frequency effect for long compounds in the first fixation duration was indeed attenuated in Experiment 2, we pooled the data for the long compounds of both experiments and used experiment as a between-participant and between-item variable. If the frequency effect were indeed attenuated, there should be an interaction between experiment and frequency. The ANOVAs showed a significant effect of frequency $F1(1,58) = 29.4$, $p < .001$; $F2(1,38) = 8.5$, $p < .01$) and a tendency for the Frequency × Experiment interaction ($F1(1,58) = 11.1$, $p < .005$; $F2(1,38) = 3.2$, $p = .08$). Since the t tests for long compounds in Experiment 1 were highly significant, we were inspired to run the t tests on the long compounds of Experiment 2 as well. The separate t tests showed that the frequency effect for long compounds in Experiment 2 was at best only marginal in the participant analysis ($t1(29) = 1.7$, $p = .097$; $t2(38) = 1.2$, $p = .239$).[11]

The question then arises: Does whole-word frequency affect first fixation durations of long compounds? The answer, to our mind, is 'possibly.' The main effect of frequency in Experiment 2 and the tendency for a whole-word frequency effect ($p1 = .05$, $p2 = .11$) in Pollatsek et al. (2000) indicate that the whole-word form might be involved in the early stages of processing long compounds. However, the small effect sizes in both Pollatsek et al. (5 ms) and here in Experiment 2 (4 ms), combined with the tendency for an interaction between experiment and frequency observed in this study, suggest that, if anything, the involvement is rather small.

Discussion

Frequency effects. When manipulating the whole-word frequency for long and short compounds, the pattern of results appears quite different from that obtained in Experiment 1. With respect to the measures that reflect early processing stages, we find, if anything, a weak effect for long compounds in first fixation duration, and no effect in intraword saccade length. The latter indicates that, at the early stages, there is no difference in parafoveal intraword preprocessing of high-frequency and low-frequency long compounds. The

[10] At first sight, one might think that the difference in effect size is caused by a difference in frequency manipulation between Experiments 1 and 2 with a stronger frequency manipulation for Experiment 1. In Experiment 1, the frequency manipulation is 472 vs. 23 (a ratio of 20.5:1), whereas in Experiment 2 the manipulation is 25 vs. 2.3 (a ratio of 10.9:1). However, the reported frequency values here are absolute values and it is well known that normally response latencies or reading times are in an inverse relationship with the logarithmic transformation of frequency values. The frequency manipulation in Experiment 1 on the basis of average log frequencies is 2.56 vs. 1.11 (ratio 2.3:1), in Experiment 2 it is 1.34 against 0.46 (2.9:1). Thus, if anything, the frequency manipulation of Experiment 2 was in effect stronger than that of Experiment 1.

[11] This indicates that the frequency effect in Experiment 2 for short compounds must be a solid one. Separate t test showed that this was indeed the case ($t1(29) = 2.7$, $p = .01$; $t2(38) = 2.3$, $p = .03$).

former indicates that the whole-word form might be involved in foveal processing from the initial fixation onwards. Perhaps, as assumed in the EZ-Reader model (Reichle et al., 1998), there is a familiarity check of the whole word at the first fixation preceding complete lexical access. This could imply that some aspects of the whole-word become available somewhat faster for a long high-frequency compound than for a low-frequency one, without the whole word being identified.

In contrast to long compounds, solid frequency effects were observed for short compounds in first fixation duration. This indicates that the whole-word form becomes available immediately in processing short compounds. However, again one might start to wonder why, in short compounds, a second fixation is needed as often as it is, if the whole word can so effectively be analyzed during the first fixation. As noted earlier, this might simply be related to the accumulation of activation of the whole-word form, without having reached completion of lexical access. The differences in a measure like second fixation probability (see Footnote 7) between high- and low-frequency short compounds would then be a result of the former word types more often having reached or being closer to lexical access than the latter.

At later stages of processing, the whole-word frequency exerts reliable effects for both long and short compounds. For both types of compounds positive frequency effects were obtained for third fixation probability. The solid frequency effect for both types of compounds in gaze duration demonstrates that, overall, the whole-word form plays a relevant role in processing both long and short compounds.

A particularly interesting aspect of the data of the current study is the lack of a frequency effect for saccade length in Experiment 2 for long compounds, in contrast to the positive frequency effect of this measure in Experiment 1. These results compare favorably with those obtained by Hyönä and Pollatsek (1998) and Pollatsek et al. (2000). In other words, the length of the first within-word saccade in long compounds is affected by the frequency of the first constituent but not by the frequency of the whole compound. How could these results be explained? One possible explanation is in accordance with the foveal processing difficulty hypothesis proposed by Henderson and Ferreira (1990) (see also Hyönä, 1995 and Hyönä & Pollatsek, 2000), which holds that frequency characteristics of the word in foveal vision affect the amount of information that can be analyzed or preprocessed in the parafovea. Applying this phenomenon to processes within the long compound words, it can be hypothesized that more resources can be allocated to parafoveal processes in the case of high-frequency first constituents than in the case of low-frequency first constituents, leading to effectively longer saccades in the former case. When the whole-word frequency is manipulated, early foveal efforts seem to be practically equal for both cases, and therefore parafoveal processing would be equally effective, leading to similar saccade lengths in the high-frequency and low-frequency instances. In other words, the saccade length effect for long compounds in Experiment 1 but not in Experiment 2 is taken as further support for the immediate influence of the first-constituent representation, and for the delayed impact of the whole-word form.

Word length effects. As in Experiment 1, long compounds were gazed upon longer than short compounds and they elicited more second, third and fourth fixations. Thus compounds of 8 characters or less were usually dealt with in one fixation, while compounds of 12 characters or more typically elicited two or more fixations. Therefore, one could assume that on the basis of visual acuity constraints, whole-word access is more likely for short than for long compounds.

Perhaps more surprising was the shorter first fixations for long than for short compounds in both experiments. However, in the eye movement literature, it is a well-established fact that an increase in the number of eye fixations goes hand in hand with a decrease in individual fixation times (see, e.g., Hyönä, 1995; Kliegl, Olson, & Davidson, 1983; O'Regan & Lévy-Schoen, 1987). Note that this state of affairs yielded a better chance for an early first-constituent frequency effect for short than for long compounds, since the time-span during which a possible first-constituent frequency effect could have occurred was longer for short than for long compounds.[12]

General discussion

The question that we have addressed in this paper is whether the role of morphology is modified by word length. The results of the experiments showed that this question has to be answered in the affirmative. Experiment 1 revealed a first-constituent frequency effect for long compounds in several measures: gaze duration, first fixation duration, saccade length, and third fixation probability. More precisely, long compound words with a high-frequency first constituent elicited shorter gaze durations, shorter first fixation durations, longer intra-word saccades from first to second fixation location, and

[12] Another common phenomenon in eye movement behavior is that the first fixation is located further into long than short words as was observed in both experiments of this study (see also Inhoff, 1989; McConkie, Kerr, Reddix, & Zola, 1988; O'Regan, 1992; Rayner, 1979). This is a result of parafoveal preprocessing of the target word on $N-1$, during which low-level features of the upcoming word like word length are picked up. Accordingly, a saccade is programmed to the so-called preferred landing position that is somewhat to the left of the word center.

fewer third fixations than long compounds with a low-frequency first constituent. In contrast, there were no solid effects of the first-constituent frequency manipulation in any of the measures for short compounds.

In Experiment 2, the situation was quite different. Most importantly, only mild, if any, whole-word frequency effects were found for long compound words in first fixation duration, and no effect in intraword saccade length. However, in two late measures, the whole-word frequency manipulation did exert an effect. More specifically, long compound words with high whole-word frequency elicited shorter gaze durations and fewer third fixations than long compounds words with low whole-word frequency. For the short compounds, the picture changed more radically: the whole-word frequency manipulation now elicited significant effects in several measures. Short compounds with high whole-word frequency elicited shorter gaze durations, first fixation durations and less third fixations than short compounds with low whole-word frequency. Table 5 summarizes all the frequency effects for the two word types in Experiments 1 and 2.

The table is divided into four subsections reflecting different stages of processing. The preprocessing measure considered here does not show any frequency effect. The pattern of results for measures that reflect early stages of processing reveals remarkable differences between short and long compounds. The most marked difference is that the frequency effect observed in first fixation duration is triggered by a different kind of frequency manipulation: it is mainly triggered by a first-constituent manipulation for long compounds and only by a whole-word frequency manipulation for short compounds. This strongly suggests that the whole-word form of short compounds is involved in processing from the first fixation onwards, whereas for long compounds the initial phase is mainly restricted to the processing of the first constituent. The latter conclusion is also supported by the intraword saccade length data. Third fixation probability, a measure that reflects a late processing stage, indicates that first-constituent frequency effects linger on for long compounds and do not notably emerge for short compounds. On the other hand, whole-word frequency effects for this late measure are solid and stable for both types of compounds. Thus, it seems, that for long compounds the first constituent is processed first, and the whole-word form gets involved in the process at a later stage. For short compounds, the whole-word form continues to be involved at a later stage. Finally, gaze duration (a summary measure of fixation durations and fixation probabilities) is similarly affected by the whole-word frequency manipulation for both types of compounds, but differently by the first-constituent frequency manipulation. The latter had no effect on short compounds, indicating once more that, overall, this type of compound is processed via the whole-word form.

Morphology and word length

As observed in the present and several previous studies, morphological structure can affect the way complex words are processed. For compound words in particular, the majority of studies shows that, at some level of processing, constituent morphemes are involved in lexical processing. In most theoretical accounts, this involvement is thought to take place at the early stages of processing (Jarema et al., 1999; Sandra, 1990; Zwitserlood, 1994). The question that we asked is whether this early involvement is guided by structural principles or by visual principles. If the former were the case, it should not depend on the length of the compound whether or not early effects of morphological structure are observed. In the current study, we showed that this is not the case. For long compounds, early effects of morphological structure were abundantly present, whereas for short compounds with exactly the same structure and exactly the same lexical–statistical characteristics no such effects were found. On the basis of these observations, it can be concluded that visual constraints modify the role of morphology in processing compound words. More precisely, it seems that the first constituent of a long compound has what we have called a visual acuity benefit over the second constituent and

Table 5
Summary of positive frequency effects for short and long compounds in Experiment 1 (first-constituent frequency manipulation) and Experiment 2 (whole-word frequency manipulation) on the following measures: initial landing position (ILP), first fixation duration (FFD), intraword saccade length (SL), probability of a third fixation (%3Fix), and gaze duration (Gaze)

Exp. + word type	Frequency manipulation	Preprocessing	Early stages of processing	Late stages of processing	Sum	
		ILP	FFD	SL	%3Fix	Gaze
1: Long compounds	First constituent	No	Yes	Yes	Yes	Yes
1: Short compounds	First constituent	No	No	–	No	No
2: Long compounds	Whole word	No	Mildly	No	Yes	Yes
2: Short compounds	Whole word	No	Yes	–	Yes	Yes

the whole word. It is purely for this reason that processing of the first constituent is initiated before that of the second constituent and the whole word. One might argue that the whole-word form starts to affect the processing of long compounds from the initial fixation onwards as well, but according to our and the Pollatsek et al. (2000) data, it does so only to a small extent if at all.

When all the sublexical and lexical information can be extracted during a single fixation, as is the case for the short compounds, whole-word information seems to become available earlier than first-constituent information. The system seems to be tuned in such a way that it starts to analyze the string of letters that is under foveal inspection. If this string is around 8 letters (or less) it starts to analyze the whole word, whereas if the letter string is longer, it will start to analyze the beginning of the word to acquire access to a meaningful substring. Note that in the case of a short compound, information on the morphological substrings is, in principle, available right from the beginning and could, therefore, have affected processing at the early stages. In the case of a long compound, analysis could have been postponed until all characters were foveally inspected to extract whole-word information. However, our data do not support either possibility. Morphological effects tend to be more pronounced and kick in earlier when compound words are long and therefore typically elicit multiple fixations, than when they are short and as a rule elicit one fixation only.

One might argue that the observed decomposition effects are easier to obtain in languages in which compounding is abundantly present like Finnish and German. However, a whole body of evidence contradicts this; it has been observed that decomposition is a fairly common phenomenon for many types of languages, from Hebrew to Dutch, from Italian to Chinese, and from English to Finnish. There are two recent studies of English compounds that failed to find a systematic effect of first-constituent frequency in the first fixation duration. We would like to argue that the lack of clearly significant early effects is due to their compounds being relatively short. Juhasz et al. (in press) found no or mild first-constituent frequency effects for English compounds of nine letters in lexical decision, naming, and eye movement measures such as first fixation duration and gaze duration. The tendencies for such an effect were slightly stronger than in our study, perhaps in accordance with the fact that their compound words were slightly longer than our short compounds. However, apparently they were still not long enough to exert reliable effects in participants and items analyses at the 5% level. Similarly, Andrews et al. (in press) failed to find a solid, early effect of first-constituent frequency for English compound words with an average length of about 8.5 characters. Instead, regression analyses of both their experiments showed that whole-word frequency is the strongest predictor of first fixation duration and gaze duration. In agreement with the main point of the present study, the authors note that 'early constituents of long words may play a dominant role in processing simply because this is where people tend to land their eyes, rather than because the first segment has an inherently special role in the retrieval process.'

Word length seems to be a modifying factor not only for compound words, but also for derived words and inflected words, as can be gathered from the studies of Colé et al. (1989), Beauvillain (1996), and Niswander et al. (2000). Thus, in French, Colé et al. (1989) observed whole-word frequency effects only for rather short prefixed words, but Beauvillain (1996) observed solid base frequency effects when employing prefixed words of 10 characters or more. Niswander et al. (2000) found that also in English morphological effects come and go if you extend or shorten the length of complex words other than compounds. More precisely, they found solid base frequency effects for long derived words, and only surface frequency effects for short inflected words.

As noted before, these studies were not designed to directly test the modulating effect of word length, whereas the present study was. For the long and short conditions the same word types were used, the frequency manipulations were identical and other factors were matched in an identical way. Even under these circumstances, word length exerted a pervasive impact on the involvement of sublexical units in word processing. In addition, the present study shows that the modulating role of word length is not restricted to laboratory tasks alone (as used in the French studies), but that it can also be observed in a normal reading paradigm as employed here.

Implications for morphological processing models

On a very general level, one could say that all models that assign some kind of role to morphology in processing complex words are correct (e.g., Sandra, 1990, Zwitserlood, 1994) However, not many models could be rejected on the basis of this notion, for even in whole-word accounts, morphological relations between complex words are often incorporated (e.g., De Jong et al., 2001; Giraudo & Grainger, 2000). The beauty of studying morphological processing with eye movements is that one can acquire a temporal insight into the availability of certain information. If whole-word frequency and morphemic frequency effects are found in a lexical decision study, all one can say is that at some stage of processing the whole-word form and morphemic constituents are involved. It would be impossible to decide (even though researchers often do) whether access is sublexical (from morphemic to lexical level), supralexical (from lexical to morphemic level) or simultaneous. However, on the basis of the current eye movement study, it can be concluded that models that assume sublexical access of complex words like Taft and Forster's (1976) search model or Taft's (1994) interactive model cannot account for the early effects of whole-word frequency as observed for the short compounds. On the other hand, early effects of first-constituent frequency as observed for the long compounds in the present and earlier studies (e.g., Hyönä & Pollatsek, 1998) cannot be explained by models that assume supralexical access only at an early stage of processing complex words (e.g., De Jong et al., 2001; Giraudo & Grainger, 2000; Van Jaarsveld & Rattink, 1988; Zhou & Marslen-Wilson, 2000). However, it should be noted that early effects of first-constituent frequency are in line with models that assign a prominent role to the first constituent (Mäkisalo et al., 1999; Taft & Forster, 1976). Race models that assume a parallel horse race right from the start between the direct and the decomposition route (Baayen et al., 1997; Bertram et al., 2000a,b,c; Libben, 1993; Schreuder & Baayen, 1995) would predict similar effects of whole-word and first-constituent frequency for long and short compounds, which is at odds with our data. Thus, it seems that the differential pattern observed for long and short compounds is not easily accounted for by any model.

The next question to be asked is which model would fare best if only short or long compounds are considered. Even an answer to this question is not straightforward. For short compounds, we can only consider models that postulate initial access via the whole-word form. The supralexical model of Giraudo and Grainger (2000) makes this postulation and is therefore in line with the early whole-word frequency effects as found in Experiment 2. However, the lack of morphological effects at a later stage (as reflected in third fixation probability, but also in gaze duration to some extent) is problematic for the model as it holds that, after whole-word access, the morphemic constituents will be activated as well. A race model like the one Schreuder and Baayen (1995) propose would hold that early whole-word effects are the result of a horse race in which the direct route and decomposition route compete with each other, and in which the former would usually be first. Later morphological effects can appear, but only when the whole-word recognition process runs into problems. With the reasonably familiar short compounds, this is not very likely. This type of race model would therefore fare rather well with short Finnish compounds. With long compounds, the situation is more complicated. The early effect of first-constituent frequency seems to call for a decomposition-type model. The model of Taft and Forster (1976), which claims that the first constituent serves as a lexical entry, seems a fairly good candidate. However, since the influence of the first constituent lingers on for long compounds (as can be derived from the frequency effect in third fixation probability and in gaze duration) and since the influence of whole-word frequency becomes apparent from the second and possibly first fixation onwards, one is forced to argue that at least at some point in time both sublexical and lexical information are activated simultaneously. This is evidence against the Taft and Forster (1976) model or against any model that is strictly serial in nature. Race models would explain the strong first-constituent frequency effect at early stages by an initially faster decomposition than direct route, although some kind of familiarization process with the whole-word form might also be initiated from the first fixation onwards. The late effects of whole-word frequency would indicate that, at a later stage, the direct route is also effectively involved in the recognition process. In the morphological race framework, this might well be a consequence of the fact that the decomposition route is a slow one to complete, giving the direct route a chance to become strongly involved in the recognition process. In order to complete the decomposition route, access to the second constituent is needed as well. As we know from our previous study (Pollatsek et al., 2000) this process is only initiated from the second fixation onwards. After access of the second constituent, a compositional process is required to combine the information from the first and second constituents. The strong involvement of the direct route could be a consequence of the extra operations required for the decomposition route. This also implies that information from the first constituent has to be kept active for later stages of processing, which is reflected in the first-constituent frequency effects observed in late measures (probability of a third fixation, gaze duration).

In the race model discussed above, decomposed and whole-word access are thought to operate independently from each other. In the interactive activation model of Taft (1994) decomposed and whole-word access is a more cooperative process. In fact, there is only one route that, as explained above, is sublexical, in that it passes through a morphemic level before it reaches a lexical level. However, activation at lower levels works its way up to higher levels and activation feedback from these higher levels can enhance the activation of promising lower-level units. It is easy to see that this kind of model would predict early first-constituent frequency effects, but also later whole-word frequency effects and, by virtue of activation feedback, later first-constituent frequency effects. This model also seems to imply that information from the first constituent is fed forward to the lexical level, where several candidates, starting with this very constituent, are activated (cf. the cohort model for spoken word recognition of Marslen-Wilson, 1987). The strength of activation is thought to be a function of frequency, with more activation for a high-frequency than for a low-frequency word. Another possibility would be that information from the second constituent is required, before whole-word candidates at the lexical level are activated. If this were the case, one would expect second-constituent frequency effects to appear before whole-word frequency effects. However, Pollatsek et al. (2000) found evidence that effects of whole-word frequency are at least as early as those of second-constituent frequency. In other words, Taft's (1994) model, in which initial access goes via the first constituent, after which several candidates are, depending on their frequency, more or less activated, accounts very well for the pattern of results we obtain for long compounds.

Thus, interactive activation models seem to account for specific domains of compound processing. An interactive supralexical model accounts for the early effects we obtain with short compounds, and an interactive sublexical model accounts well for the effects we obtain with long compounds. Race models can in principle account for the pattern of results of both types of compound when considered separately. However, they cannot account at the same time for both of them since all the frequency values are exactly the same for long and short compounds. In other words, if one route dominated over the other in one type of compound, it should have done so also in the other type of compound. This was clearly not the case, so we have to conclude that none of the existing race models accounts for the full pattern of results. If, however, a race model like the one proposed by Schreuder and Baayen (1995) incorporated a headstart in the case of long compounds due to a visual acuity benefit of the first constituent, the results obtained here could be accounted for. In other words, this updated model would state that the processing system is geared to acquire as much information during a single fixation as possible. This means that in the case of short compounds, both direct route and decomposition route can be initiated immediately, whereas in the case of long compounds, the direct route has to be more or less delayed until the second fixation when all the characters of a word have been foveally inspected.

In sum, word length is a factor that should be added to the list of factors that affect morphological processing. Our results show a consistently different pattern for long and short compounds. It is notable that none of the current models of morphological processing can account for the pattern of results obtained in this study. It should also be stressed once more that the employment of eye-tracking allowed us to obtain a good temporal insight into the role of morphological structure in lexical processing, something that cannot be achieved by most of the methods that are currently employed to study morphological processing. More specifically, the eye movement paradigm allows the testing of more detailed hypotheses of current models of morphological processing. Therefore more eye-tracking studies on morphological processing are called for.

References

Andrews, S. (1986). Morphological influences on lexical access: Lexical or nonlexical effects. *Journal of Memory and Language, 25*, 726–740.

Andrews, S., Miller, B., & Rayner, K. (in press). Eye movements and morphological segmentation of compound words: There is a mouse in the mousetrap. *European Journal of Cognitive Psychology*.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual route model. *Journal of Memory and Language, 36*, 94–117.

Beauvillain, C. (1996). The integration of morphological and whole-word form information during eye fixations on prefixed and suffixed words. *Journal of Memory and Language, 35*, 801–820.

Bertram, R., Laine, M., & Karvinen, K. (1999). The interplay of word formation type, affixal homonymy, and productivity in lexical processing: Evidence from a morphologically rich language. *Journal of Psycholinguistic Research, 28*, 213–226.

Bertram, R., Hyönä, J., & Laine, M. (2000a). The role of context in morphological processing: Evidence from Finnish. *Language and Cognitive Processes, 15*, 367–388.

Bertram, R., Laine, M., Baayen, R. H., Schreuder, R., & Hyönä, J. (2000b). Affixal homonymy triggers full-form storage, even with inflected words, even in a morphologically rich language. *Cognition, 74*, B13–B25.

Bertram, R., Schreuder, R., & Baayen, R. H. (2000c). The balance of storage and computation in morphological processing: The role of word formation type, affixal homonymy and productivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 489–511.

Bradley, D. C. (1980). Lexical representation of derivational relation. In M. Aronoff & M.-L. Kean (Eds.), *Juncture* (pp. 37–55). Saratoga, CA: Anma Libri.

Burani, C., & Caramazza, A. (1987). Representation and processing of derived words. *Language and Cognitive Processes, 2*, 217–227.

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Language production (Vol. II): Development, writing and other language processes* (pp. 257–294). London: Academic Press.

Colé, P., Beauvillain, C., & Segui, J. (1989). On the representation and processing of prefixed and suffixed derived words: A differential frequency effect. *Journal of Memory and Language, 28*, 1–13.

De Jong, N. H., Feldman, L. B., Schreuder, R., Pastizzo, P., & Baayen, R. H. (2001). The processing and representation of Dutch and English compounds: Peripheral morphological and central orthographic effects. *Brain and Language, 73*, 1–13.

Giraudo, H., & Grainger, J. (2000). Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language and Cognitive Processes, 15*, 421–444.

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.

Henderson, J. M., & Ferreira, F. (1993). Eye movement control during reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology, 47*, 201–221.

Hittmair-Delazer, M., Andree, B, Semenza, C., & De Bleser, R. (1994). Naming by German compounds. *Journal of Neurolinguistics, 8*, 27–41.

Hyönä, J. (1995). Do irregular letter combinations attract readers' attention. Evidence from fixation locations in words. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 68–81.

Hyönä, J., & Olson, R. K. (1995). Eye movement patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1430–1440.

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception & Performance, 24*, 1612–1627.

Hyönä, J., & Pollatsek, A. (2000). Processing of Finnish compound words in reading. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 65–87). Amsterdam: Elsevier.

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., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics, 40*, 431–439.

Inhoff, A. W., Radach, R., & Heller, D. (2000). Complex compounds in German: Interword spaces facilitates segmentation but hinder assignment of meaning. *Journal of Memory and Language, 42*, 23–50.

Jarema, G., Busson, C., Nikolova, R., Tsapkini, K., & Libben, G. (1999). Processing compounds: A cross-linguistic study. *Brain and Language, 68*, 362–369.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 824–843.

Juhasz, B. J., Starr, M. S., Inhoff, A. W., & Placke, L. (in press). The effects of morphology on the processing of compound words: Evidence from naming, lexical decisions and eye fixations. *British Journal of Psychology*.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehending. *Psychological Review, 87*, 329–354.

Kliegl, R., Olson, R. K., & Davidson, B. J. (1983). On problems of unconfounding perceptual and language processes. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 333–343). New York: Academic Press.

Laine, M., & Virtanen, P. (1999). WordMill Lexical Search Program. Center for Cognitive Neuroscience, University of Turku, Finland.

Laudanna, A., & Burani, C. (1995). Distributional properties of derivational affixes: Implications for processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 345–364). Hillsdale, NJ: Erlbaum.

Levy-Schoen, A., & O'Regan, K. (1979). The control of eye movements in reading. In P. A. Kolers, M. E. Wrolstad, & H. Bouma (Eds.), *Processing of visible language II* (pp. 7–36). New York: Plenum Press.

Libben, G. (1993). A case of obligatory access to morphological constituents. *Nordic Journal of Linguistics, 16*, 111–121.

Libben, G. (1998). Semantic transparency in the processing of compounds: Consequences for representation, processing, and impairment. *Brain and Language, 61*, 30–44.

Lima, S. D., & Pollatsek, A. (1983). Lexical access via an orthographic code? The basic orthographic syllable structure (BOSS) reconsidered. *Journal of Verbal Learning and Verbal Behavior, 22*, 310–322.

Mäkisalo, J., Niemi, J., & Laine, M. (1999). Finnish compound structure: Experiments with a morphologically impaired patient. *Brain and Language, 68*, 249–253.

Marslen-Wilson, W. (1987). Parallel processing in spoken word recognition. *Cognition, 25*, 71–102.

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: The location of initial fixation on words. *Vision Research, 28*, 1107–1118.

McQueen, J. M., & Cutler, A. (1998). Morphology in word recognition. In A. Zwicky & A. Spencer (Eds.), *The handbook of morphology*. Oxford: Blackwell.

Monsell, S. (1984). Repetition and the lexicon. In A. W. Ellis (Ed.), *Progress in the psychology of language* (2), (pp. 147–196). Hove: Erlbaum.

Niswander, E., Pollatsek, A., & Rayner, K. (2000). The processing of derived and inflected suffixed words during reading. *Language and Cognitive Processes, 15*, 389–420.

O'Regan, J. K. (1992). Optimal viewing position in words and the strategy-tactics theory of eye movements in reading. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 334–354). New York: Springer-Verlag.

O'Regan, J. K., & Lévy-Schoen, A. (1987). Eye movement strategy and tactics in word recognition and reading. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 363–383). Hillsdale, NJ: Erlbaum.

Plaut, D., & Gonnerman, L. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes, 15*, 445–486.

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.

Raney, G. E., & Rayner, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology, 49*, 151–172.

Raveh, M., & Rueckl, J. G. (2000). Equivalent effects of inflected and derived primes: Long-term morphological priming in fragment completion and lexical decision. *Journal of Memory and Language, 42*, 103–119.

Rayner, K. (1979). Eye-guidance in reading: Fixation location 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., & 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. (1976). What guides a reader's eye movements? *Vision Research, 16*, 829–837.

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 1188–1200.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*, 125–157.

Rubenstein, H., & Pollack, I. (1963). Word predictability and intelligibility. *Journal of Verbal Language and Verbal Behavior, 2*, 147–158.

Sandra, D. (1990). On the representation and processing of compound words: Automatic access to constituent morphemes does not occur. *Quarterly Journal of Experimental Psychology, 42A*, 529–567.

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 131–154). Hillsdale, NJ: Erlbaum.

Semenza, C., Luzzatti, C., & Carabelli, S. (1997). Morphological representation of compound nouns: A study on Italian aphasic patients. *Journal of Neurolinguistics, 10*, 33–43.

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory & Cognition, 7*, 263–272.

Taft, M. (1994). Interactive-activation as a framework for understanding morphological processing. *Language and Cognitive Processes, 9*, 271–294.

Taft, M., & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior, 15*, 607–620.

Van Jaarsveld, H., & Rattink, G. (1988). Frequency effects in the processing of lexicalized and novel nominal compounds. *Journal of Psycholinguistic Research, 17*, 447–473.

Vannest, J., & Boland, J. (1999). Lexical morphology and lexical access. *Brain and Language, 68*, 324–332.

Whaley, C. P. (1978). Word–nonword classification time. *Journal of Verbal Language and Verbal Behavior, 17*, 143–154.

Zhou, X., & Marslen-Wilson, W. D. (2000). Lexical representation of compound words: Cross-linguistic evidence. *Psychologia, 43*, 47–66.

Zwitserlood, P. (1994). The role of semantic transparency in the processing and representation of Dutch compounds. *Language and Cognitive Processes, 9*, 341–368.