

Are long compound words identified serially via their constituents? Evidence from an eye-movement–contingent display change study

JUKKA HYÖNÄ and RAYMOND BERTRAM
University of Turku, Turku, Finland

and

ALEXANDER POLLATSEK
University of Massachusetts, Amherst, Massachusetts

The processing of two-constituent 12- to 18-letter Finnish compound nouns was studied by using an eye-movement–contingent display change technique. In the display change condition, all but the first 2 letters of the second constituent were replaced by visually similar letters until the eyes moved across an invisible boundary. When the eyes crossed the boundary, the second constituent was changed to its intended form. In the control condition, there was no display change. The frequency of the first constituent was also varied. The major findings were that (1) fixation time on the first constituent was strongly affected by the frequency of the first constituent but was not at all affected by whether the second constituent was visible, but (2) fixation time on the word subsequent to the first constituent's having been left was strongly affected by the display change. These results are most parsimoniously explained by the serial access of the two constituents for these long compound words.

In our prior research, we examined the identification of two-noun Finnish compound words by embedding them in sentence contexts and recording readers' eye movements on these compounds while they read the sentences for comprehension. (In Finnish, compound words are never broken by interword spaces.) The results suggest that long (12- to 18-letter) compound words are processed primarily in such a way that the constituents are processed serially (from left to right). Hyönä and Pollatsek (1998) manipulated the frequency of the first constituent (i.e., its frequency as a separate word) while matching for second-constituent and whole-word frequency (e.g., *liikenne* in *liikennevalo* = traffic light). They observed that first-constituent frequency had both an early effect that surfaced in the duration of the initial fixation on the word and a later, larger effect on the gaze duration on a word (the total *first pass time* on the word), which was due largely to differences in the number of times that the word was fixated. Both of these effects were subsequently replicated by Bertram and Hyönä (2003) for long compounds. Pollatsek, Hyönä, and Bertram (2000) also observed that the frequency of the second constituent had a large effect on the gaze duration on the compound

word. However, second-constituent frequency did not affect the duration of first fixation on the compound but affected processing starting only on the second fixation. These data are consistent with the view that long compounds are identified serially via their constituents by first accessing the initial constituent, followed by the access of the second constituent (for other eye movement studies of compound word processing in reading, see Andrews, Miller, & Rayner, 2004; Inhoff, 1989a; Inhoff, Briehl, & Schwartz, 1996; Inhoff, Radach, & Heller, 2000; Juhasz, Starr, Inhoff, & Placke, 2003). However, since Pollatsek and Hyönä (in press) demonstrated that both the gaze durations and the first-constituent frequency effects were the same for semantically opaque compounds as for semantically transparent compounds, it appears that the initial stages of the composition of the two constituents does not involve integrating their meanings.

There are data, however, indicating that accessing a compound word is more than a serial identification of its components. In particular, when the frequency of the compound word was varied (while equating the frequency of the constituents), there were also large effects on the gaze duration on the compound word (Pollatsek et al., 2000), although the effect of word frequency was not significant until the second fixation. This word frequency effect indicates that more holistic processing of the compound words is occurring as well. Moreover, Bertram and Hyönä (2003) found that a compositional mode of processing may be important only for relatively long compound words. For shorter compound words (7–9 characters), they obtained an early effect of word frequency but

This research was supported by an Academy of Finland grant to the first author and by Grant HD26765 from the National Institutes of Health to the third author. We are grateful to Annu Haapakangas and Tuomo Häikiö for their help in collecting and analyzing the data. Correspondence should be addressed to J. Hyönä, Department of Psychology, University of Turku, Turku FIN-20014, Finland (e-mail: hyona@utu.fi).

no clear effects of first constituent frequency. They argued that visual acuity constraints modulate the extent to which compounds are processed serially via their constituents or in parallel via their whole-word form. That is, when all letters fall in foveal vision (as is the case with short compounds), the whole-word form may be directly accessed, whereas when the last letters to the right of fixation remain outside foveal vision, processing is more likely to start off with the access of the first constituent, followed by the access of the second constituent and the whole word (the latter two access processes presumably overlap in time).

The privileged status of the first constituent in early stages of processing has not come out as strongly for English compounds. Andrews et al. (2004) and Juhasz et al. (2003) have found only marginally significant effects of first-constituent frequency on first-fixation duration. It should be noted, however, that both studies examined compounds that were shorter in length than those in our Finnish experiments (Andrews et al. used 6- to 11-letter compounds, and Juhasz et al. used 9-letter compounds). As was mentioned above, Bertram and Hyönä (2003) found no effect of first-constituent frequency on the first-fixation duration for shorter (7- to 9-letter) compounds, whereas the effect was clearly reliable for longer compounds. Thus, we think that the most parsimonious explanation for the data is that the role of the first constituent in compound word processing is mediated by the length of the compound word and that the somewhat different results in the two languages do not require any deeper explanation. Finally, it is worth noting that in both English studies, significant (or nearly significant) effects of second constituent frequency on gaze duration were observed—an effect also observed for Finnish (Pollatsek et al., 2000).

The present experiment was designed as a further test of the serial componential view proposed above for the processing of long compound words. In the experiment reported below, we employed a new method, which is a modification of the boundary technique developed by Rayner (1975). The technique was originally developed to study the types of information extracted from parafoveally presented words prior to when they are fixated. A typical experimental setup is as follows. The parafoveal target word is initially changed by replacing it with, for example, a semantically or orthographically similar word. An invisible boundary is placed to the left of the parafoveal word, and when the eyes cross the boundary, the parafoveal word is changed to its intended form. The display change is made during the saccadic eye movement, and since vision is greatly suppressed during saccades, the reader does not perceive the actual change taking place.

A key finding is that the subsequent foveal processing of the word benefits from a parafoveal preview of the word (i.e., the so-called parafoveal preview benefit effect; see Rayner, 1998, for a review). The parafoveal processing benefit appears to be limited mostly to ortho-

graphic and phonological information, and the bulk of the evidence suggests that lexical or semantic information is not used as preview information (for a review of the mostly negative effects, see Rayner, White, Kambe, Miller, & Liversedge, 2003). In Hebrew, parafoveal processing is extended to morphological properties (Deutsch, Frost, Pelleg, Pollatsek, & Rayner, 2003), which does not appear to be the case in English (Inhoff, 1989a; Lima, 1987). Finally, most studies suggest that the preview benefit is not “cashed in” until the parafoveal word is fixated; in other words, parafoveal processing is not found to influence foveal processing (but see Kennedy, Pynte, & Ducrot, 2002, for parafoveal-on-foveal effects).

In the present experiment, we introduced display changes to the second constituents of long compound words in order to examine directly the extent to which the second constituent is processed while the first constituent is fixated on and the effect this processing has (if any) on these initial fixations on the first constituent. To our knowledge, this is the first time the boundary technique has been employed to study within-word parafoveal processing. If a serial view of compound word processing is correct, changes made to the second constituent should not affect the processing of the first constituent. On the other hand, if the reader attempts to access the whole word form at once and/or information from the second constituent influences the processing of the first constituent, there may be an effect of change observable even when the first constituent is still fixated. The invisible boundary was set at the constituent boundary. In the display change condition, we replaced the last letters of the second constituent with visually similar letters, but the first two letters of the second constituent were always kept intact. The display change condition was then compared with a no-change condition.

First-constituent frequency was also manipulated. Thus, we compared a group of compounds that had a low-frequency initial constituent with another group of compounds with a high-frequency initial constituent. The primary motivation for the frequency manipulation was to have a “marker” for the processing of the first constituent. That is, as was indicated earlier, Hyönä and Pollatsek (1998) found that the frequency of the first constituent affected processing even on the first fixation. Thus, looking at the time course of the first fixation frequency effects should give us a good measure of when the first constituent is being processed. A second motivation for manipulating the frequency of the first constituent was to test the view that processing load is capable of affecting perceptual span (Bertram & Hyönä, 2003; Henderson & Ferreira, 1990; Hyönä & Pollatsek, 2000). On the assumption that the processing load is increased when a low-frequency constituent is fixated, the perceptual span may be assumed to constrict, which in turn would cause the display change effect to be smaller than that in the high-frequency condition. In addition, if the perceptual span is larger when a high-frequency constituent is initially

fixed, we may observe an early effect of display change in the high-frequency condition—a finding that would contradict a strictly serial view.

METHOD

Participants

Twenty-three university students took part in the experiment for course credit. All were native speakers of Finnish.

Apparatus

Eye movements were collected by the EYELINK II eyetracker manufactured by SR Research Ltd. (Toronto). The eyetracker is an infrared video-based tracking system combined with hyperacuity image processing. There are two cameras (one for each eye) that are mounted on a headband including two infrared light emitting diodes (LEDs) for illuminating each eye. The apparatus (headband plus cameras) weighs 450 g. The cameras sample pupil location and pupil size at the rate of 500 Hz. Registration is monocular and is performed for the selected eye by placing the camera and the infrared light source 4–6 cm away from the eye. Spatial accuracy is better than 0.5°. Head position with respect to the computer screen is monitored with the help of a head-tracking camera mounted on the center of the headband at the level of the forehead. Four LEDs, attached to the corners of the computer screen, are viewed by the head-tracking camera once the participant sits directly facing the screen. Head motion is detected as movements of the four LEDs and is compensated for on line from the eye position records.

Materials

A set of 48 low-frequency first constituent compounds were paired with a set of 48 high-frequency first constituent compounds. All the target words were two-noun compound words (12–18 characters long). The constituent frequency refers to the frequency the constituent has as a separate word in Finnish. The frequencies were computed on the basis of an unpublished 22.7 million word newspaper corpus of Laine and Virtanen (1999). As can be seen in Table 1, the two word sets differed substantially in the frequency of the first constituent but were matched for whole-word frequency, word length, second-constituent frequency, and the lengths of the constituents.

The target words appeared near the beginning of sentences, but never in the initial word position. Each target word with an infrequent first constituent was paired with one that had a frequent initial constituent, and a sentence frame was created for this word pair that was identical up through the word following the target word. A typical sentence pair is as follows, with the target word in bold (a slash marks the constituent boundary; no slash was present in the experimental items).

Low-frequency first constituent condition:

No Change: Käsittäakseni **vanilja/kastike** voi palaa pohjaan melko helposti.

Change: Käsittäakseni **vanilja/kaeflha** voi palaa pohjaan melko helposti.

“As far as I know **vanilla sauce** can burn pretty easily.”

High-frequency first constituent condition:

No Change: Käsittäakseni **tuotanto/volyymi** voi kasvaa tänä vuonna huomattavasti.

Change: Käsittäakseni **tuotanto/voignnl** voi kasvaa tänä vuonna huomattavasti.

“As far as I know **production volume** can grow significantly this year.”

The target sentences were presented in Courier font (so that each character position was of equal width) starting from near the left of the computer screen. The sentences occupied a maximum of two lines of text, and the critical word always appeared on the first line but was never the initial or the final word of the line. With a viewing distance of about 60 cm, each character subtended approximately 0.3° of visual angle. There were four blocks of sentences; each member of the matched quadruplet appeared in a separate block. However, each participant saw only two of the four blocks, thereby seeing all the target words only once (i.e., in either the change or the no-change condition). There were 24 items in each condition per participant. The order of the blocks was counterbalanced across participants, and the order of the target sentences was randomized within each block. There were 80 filler sentences (consisting of the materials of another experiment) among the critical sentences.

An invisible boundary was set at the morpheme boundary. In the display change condition, when the eyes crossed this boundary, the word was changed to its correct form during the saccade across this boundary. A crucial design feature of the experiment was that we wanted to make the display change as unobtrusive as possible. Accordingly, we experimented with various possibilities for how much of the second constituent to preserve. If none of it was preserved in the display change condition, the display change was often quite noticeable. However, when the first two letters of the second constituent were preserved, the display change was rarely, if ever, detectable. Thus, in the display change condition, the first two letters of the second constituent were kept intact before the boundary crossing, and the last letters of the second constituent were replaced by visually similar letters. (This meant that the target word initially appeared as a nonword in the display change condition.) The text was presented on a 17-in. ViewSonic (P775) monitor as white against a dark background. The refresh rate of the monitor was set at 150 Hz.

Procedure

Prior to the experiment, the eyetracker was calibrated using a 9-point calibration grid that extended over the entire computer

Table 1
Lexical-Statistical Properties for the Compound Words With a High-Frequency Versus a Low-Frequency First Constituent

Property	Compounds With High-Frequency First Constituent	Compounds With Low-Frequency First Constituent
Number of items	48	48
Mean first-constituent frequency*	344.4	8.3
Mean first-constituent length†	7.2	7.6
Mean second-constituent frequency*	173.0	188.2
Mean surface frequency*	2.2	2.0
Mean bigram frequency‡	7.2	7.7
Mean initial trigram frequency‡	0.7	0.9
Mean final trigram frequency‡	1.0	1.0
Mean word length†	13.4	13.5

*Values scaled to one million. †Word length in characters. ‡Values scaled to one thousand.

screen. Prior to each sentence, the calibration was checked by presenting a fixation point just to the left of where the beginning of the sentence would appear (at the left side of the screen, about halfway down). If needed, the calibration was automatically corrected, after which the sentence was presented.

The participants were instructed to read the sentences for comprehension at their own pace. They were told that they would periodically be asked to paraphrase the last sentence they had read, to make sure that they had attended to and understood what they had read. It was emphasized that the task was to comprehend—not to memorize—the sentences.

RESULTS

Repeated measures analyses of variance were computed on the dependent measures, with two within-subjects and within-items variables: frequency of first constituent (high vs. low) and display change (change vs. no change). In the display change condition, 19.7% of the trials were excluded from the analyses, due to the change's taking place after a fixation had already started on the second constituent (these were trials in which the eyes crossed the invisible boundary toward the end of the critical saccade).¹ We will start by reporting the gaze duration, which is the summed duration of fixations landing on the target before exiting it, followed by more detailed analyses of the eye movement pattern. The means of the eye movement measures are given in Table 2.

Gaze Duration

Gaze duration is a composite measure of fixation time spent on a target word the first time it is fixated. That is, it is the sum of the durations of all the fixations on the word on the *initial visit* before the first saccade off of the word (either to the right or the left). (Words that were skipped were excluded from the analysis.) There was a 66-msec main effect of first-constituent frequency [$F_1(1,22) = 34.07$, $MS_e = 2,941$, $p < .001$; $F_2(1,47) = 15.52$, $MS_e = 15,854$, $p < .001$] and a 101-msec main ef-

fect of display change [$F_1(1,22) = 63.76$, $MS_e = 3,627$, $p < .001$; $F_2(1,47) = 59.73$, $MS_e = 6,888$, $p < .001$]. However, the size of the display change effect was different in the two first-constituent frequency conditions—120 msec for the low-frequency condition and 80 msec for the high-frequency condition—as was indicated by a significant frequency \times display change interaction [$F_1(1,22) = 5.06$, $MS_e = 1,848$, $p < .05$; $F_2(1,47) = 4.84$, $MS_e = 4,884$, $p < .05$].

The gaze duration data thus indicate that, as in prior experiments, the frequency of the initial morpheme had a large effect on the time taken to process the compound word. They also indicate that whether or not the second constituent is visible when the initial constituent is being fixated has a large effect on processing. This clearly indicates that information from the second morpheme is being extracted before it is fixated. The nature of the interaction, however, was surprising, since we had expected there to be greater extraction of information from the second constituent when the first constituent was high frequency, analogous to the finding that there is a larger preview benefit (and presumably more information extracted from the parafovea) when a fixated word is high frequency.

Subgaze Measures

The next analyses we performed were on what we termed *subgaze* measures. That is, we wanted to divide the gaze duration on the word in the display change condition into the gaze duration before the display change was made (i.e., the initial gaze duration on the first constituent) and the gaze duration after the display change. Accordingly, for the subgaze₁ measure, we summed all the fixations that were made on the first constituent prior to the first saccade's crossing the constituent boundary; this crossing coincided with the display change in the display change condition. The trials on which the initial fixation on the word crossed the boundary were excluded

Table 2
Eye Movement Measures in the Display Change and the No-Change Conditions, Separately for Compounds With a High-Frequency Versus a Low-Frequency First Constituent

Eye Fixation Measure	High-Frequency First Constituent		Difference	Low-Frequency First Constituent		Difference
	Change	No Change		Change	No Change	
	<i>M</i>	<i>M</i>		<i>M</i>	<i>M</i>	
Gaze duration*	596	515	80	682	561	120
Subgaze on first constituent*	243	242	1	292	277	15
Subgaze after change*	410	340	70	441	350	90
Initial fixation location†	4.97	4.85	.12	4.97	5.04	-.07
First-fixation duration*	223	219	4	230	231	-1
Within-word saccade length†	4.79	4.81	-.02	3.85	3.94	-.09
Probability of refixation	.92	.89	.02	.95	.91	.05
Probability of 3 fixations	.50	.32	.18	.68	.46	.21
Probability of 4 fixations	.16	.11	.05	.31	.16	.15
Probability of returning to first constituent	.27	.16	.11	.42	.31	.11
Duration of first fixation after change*	255	229	26	235	213	23
Probability of two fixations on first constituent	.10	.12	-.02	.23	.16	.07

*In milliseconds. †In character spaces.

from the subgaze₁ analyses. For the subgaze₂ analysis, we summed all the fixations that were made on the target compound words after the constituent boundary was crossed for the first time (which coincided with the display change in the display change condition); this thus included fixations that were on both constituents. (Trials with no fixations on the target word after the display change were excluded from the subgaze₂ analyses.)

As can be seen in Table 2, there was a 42-msec main effect of the first constituent frequency on subgaze₁ [$F_1(1,22) = 47.63$, $MS_e = 872$, $p < .001$; $F_2(1,47) = 29.72$, $MS_e = 3,300$, $p < .001$], but neither the 8-msec main effect of display change nor the frequency \times display change interaction was close to being significant ($ps > .14$). The pattern of results for subgaze₂, however, was quite different. There was a 70-msec effect of display change [$F_1(1,22) = 49.83$, $MS_e = 2,985$, $p < .001$; $F_2(1,47) = 41.32$, $MS_e = 7,136$, $p < .001$] but only a 21-msec main effect of frequency in the participant analysis [$F_1(1,22) = 3.83$, $MS_e = 2,393$, $p < .10$; $F_2(1,47) = 1.49$, $MS_e = 17,463$, $p > .20$]. The frequency \times display change interaction was marginal in the item, analysis [$F_1(1,22) = 1.56$, $MS_e = 1,470$, $p > .20$; $F_2(1,47) = 3.60$, $MS_e = 3,033$, $p < .10$]. This marginal interaction mirrors the interaction in the gaze duration analysis, which indicated that the low-frequency condition was more strongly affected by display change than was the high-frequency condition.

To summarize the subgaze analyses, it appeared, to at least a reasonable approximation, that the processing of the two components of the compound word was serial and that the eye movement record was affected by them in turn. That is, when people were fixating the first constituent, their fixation time on it seemed little affected by whether what was visible in the second part of the word was the second constituent or a nonword. However, the processing time subsequent to the display change was strongly affected by whether the second constituent had been present before the boundary change. This would be consistent with a serial model of processing, such as the E-Z reader model (Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 2003), in which the decision to move the eyes forward is based on the difficulty of processing what is fixated (in this case, indexed by the frequency of the first constituent) but, between the decision to move the eyes forward and when the eyes actually move forward, there is significant extraction of parafoveal information from the text to the right (in this case, the second constituent). The extraction of meaningful information from this location (in this case, when there is no display change) speeds subsequent processing.

This story is a slight idealization of the data, since it ignores the 8-msec display change effect on subgaze₁. However, we think that our characterization of the data is justified at this point, since this 8-msec effect is fairly likely to be just a Type I error. Another possibility is that this small effect is real but is due to mistargeted sac-

ades. That is, it is likely that on a small percentage of trials, a refixation on the compound word is targeted for the second constituent but falls short and lands on the first constituent. On these trials, one would expect a significant disruption to processing, since presumably, the second constituent would be the focus of attention on this mistargeted fixation. Such an explanation would still be consistent with assuming serial processing of the constituents. A second complication is the interaction between the display change manipulation and the frequency of the first constituent. However, we will defer discussion of that until after we present the rest of the data.

Earlier Measures of Processing

As Table 2 indicates, the location of the initial fixation on the target word was basically unaffected either by the frequency of the first constituent or by the display change (all $F_s < 1.2$). In contrast, the first fixation on the word was 10 msec longer for low-frequency first constituents than for high-frequency first constituents [$F_1(1,22) = 9.04$, $MS_e = 261$, $p < .01$; $F_2(1,47) = 5.88$, $MS_e = 702$, $p < .025$]; however, the 2-msec difference between the change and the no-change conditions was not close to significant ($F_s < 1$). Similarly, the length of the saccade off the first fixation on the word was affected only by the frequency of the first constituent: The first within-word saccade was 0.91 characters longer for high-frequency first constituents than for low-frequency first constituents [$F_1(1,22) = 29.59$, $MS_e = 0.643$, $p < .001$; $F_2(1,47) = 21.66$, $MS_e = 1.681$, $p < .001$], but the 0.06 character difference between display change conditions was not close to significant ($F_s < 1$). Thus, similar to what was seen for the subgaze₁ measure, there was (1) a clear indication that the first-constituent frequency had an effect on the duration of the first fixation and where it moved subsequently but (2) no indication that the nonword second constituent in the display change condition had any effect on the eyes up through the programming of the second saccade on the word. Both of these findings (together with the finding that the first-constituent frequency had no effect on the initial fixation location) replicate those of Hyönä and Pollatsek (1998).

Later Measures of Processing

Other measures, including later fixations, mirrored the effects on gaze duration (i.e., that both variables had an effect). In particular, the probability of refixating the word was higher on each fixation when there was a display change (see Table 2). Even the probability of making at least one refixation on the target word was slightly higher when there was a display change [$F_1(1,22) = 10.78$, $MS_e = 0.0024$, $p = .01$; $F_2(1,47) = 7.24$, $MS_e = 0.0047$, $p < .01$], and it was marginally higher for low-frequency first-constituent words [$F_1(1,22) = 3.11$, $MS_e = 0.0041$, $p < .10$; $F_2(1,47) = 5.27$, $MS_e = 0.0056$, $p < .05$]. The effects were relatively small due to the refixation probability approaching ceiling (over 90% of the trials produced at least one refixation). The main effects of frequency

$[F_1(1,22) = 29.13, MS_e = 0.020, p < .001; F_2(1,47) = 28.62, MS_e = 0.045, p < .001]$ and display change $[F_1(1,22) = 65.11, MS_e = 0.014, p < .001; F_2(1,47) = 55.86, MS_e = 0.028, p < .001]$ were robust for the probability of making at least three fixations on the target, as were the frequency and display change main effects for the probability of four or more fixations $[F_1(1,22) = 21.98, MS_e = 0.012, p < .001; F_2(1,47) = 17.11, MS_e = 0.040, p < .001; F_1(1,22) = 20.10, MS_e = 0.011, p < .001; F_2(1,47) = 19.04, MS_e = 0.021, p < .001, respectively]$. In addition, on the latter measure, the frequency \times display change interaction was significant $[F_1(1,22) = 10.07, MS_e = 0.0059, p < .01; F_2(1,47) = 9.92, MS_e = 0.015, p = .01]$ since the low-frequency condition was more strongly affected by the display change than was the high-frequency condition.

To obtain a more detailed picture of the refixations, we also analyzed the probability of returning to the first constituent after at least one fixation was made on the second constituent. In fact, readers were more likely to return to the first constituent when the first constituent was low frequency $[F_1(1,22) = 27.49, MS_e = 0.018, p < .001; F_2(1,47) = 26.15, MS_e = 0.036, p < .001]$ and when the second constituent was initially replaced with a nonword $[F_1(1,22) = 21.18, MS_e = 0.013, p < .001; F_2(1,47) = 31.94, MS_e = 0.025, p < .001]$, but the interaction was clearly nonsignificant ($F < 1$).

The duration of the first fixation on the target after the eyes crossed the constituent boundary produced a slightly odd effect (see Table 2). There were significant main effects of both first-constituent frequency $[F_1(1,22) = 22.91, MS_e = 325, p < .001; F_2(1,47) = 26.40, MS_e = 620, p < .001]$ and display change $[F_1(1,22) = 32.66, MS_e = 415, p < .001; F_2(1,47) = 35.14, MS_e = 756, p < .001]$. However, the first-constituent frequency effect was “backwards,” since this fixation duration was shorter for the low-frequency condition than for the high-frequency condition. An analogous effect was observed when the total fixation time on the second constituent was analyzed. As we will argue in the final part of the Results section, the effect reflects a greater predictability of the second constituent when the first constituent is of low frequency. As regards the effect of display change, on the other hand, the duration of the first fixation after change is the earliest duration measure to show an effect. The size of the effect is similar between the two frequency conditions, as indexed by a clearly nonsignificant interaction ($F_s < 1$).

Trying to Understand the Frequency \times Display Change Interaction

As was indicated above, we were puzzled by this interaction, since we expected that, if anything, the display change would produce a greater change when the first constituent was high frequency. The reasoning was that the high-frequency first constituents would require less attention when they were fixated and, thus, allow for more processing of the second constituent before it was

fixated—and hence, there would be a bigger effect of display condition for the high-frequency first-constituent words (i.e., a bigger interference effect when this information was not present). We tested several hypotheses to explain the unexpected interaction.

Two that did not succeed very well were the following. One hypothesis was that there was more processing done on the second constituents paired with the low-frequency first constituents before crossing the boundary, because there were more fixations on the low-frequency first constituents before crossing the boundary and, thus, more opportunity for parafoveal processing. In fact, the probability of refixating the first constituent before the boundary was crossed was substantially greater in the low-frequency first-constituent condition $[F_1(1,22) = 17.54, MS_e = 0.0092, p < .001; F_2(1,47) = 19.06, MS_e = 0.019, p < .001; \text{see Table 2}]$. However, when we analyzed gaze durations for only the trials in which exactly one fixation was made on the initial constituent before crossing the boundary, these one-fixation trials had the same gaze duration pattern as the gaze durations reported in Table 2; for the low-frequency condition, the change effect was 119 msec (670 vs. 551 msec, for the change and the no-change conditions, respectively), and for the high-frequency condition, the change effect was 89 msec (602 vs. 513 msec, for the change and the no-change conditions, respectively). This indicates that something other than the greater number of fixations on the first constituent in the low-frequency condition was affecting processing of the second constituent while the first was still fixated.

A second hypothesis was that the observed interaction might be due to differential effects of the display change on the ease of parsing the compound word into its two constituents. This explanation first assumes that the word must be parsed into two constituents before lexical access of the second constituent can begin. It next assumes that, for the high-frequency first constituents, the parsing job can be accomplished largely on the basis of seeing the initial constituent and the first two letters of the second constituent, whereas more information would generally be needed from the second constituent in the low-frequency condition to accomplish the parsing job. If so, this initial job of processing the second constituent (largely accomplished when the first constituent is being fixated) would be affected greatly by the display change in the low-frequency condition but less affected by the display change in the high-frequency condition. However, since we do not have access to data that would support this parsing hypothesis, it should be considered a speculative option at this point.

The third hypothesis for the interaction is that processing of the second constituent may be easier in the low-frequency condition because the second constituent is more constrained and/or predictable than in the high-frequency condition. More specifically, we have observed in our corpus searches that, in general, fewer compound words can be formed with low-frequency than with high-frequency first constituents. Thus, pro-

cessing of the second constituent when the first is of low frequency would be easier if this smaller set of compound words does, in fact, exercise some constraint on the encoding of the second constituent. In fact, Hyönä and Pollatsek (2000) reported that fixation time on the second constituent was significantly shorter when the first constituent was of low frequency than when it was of high frequency. If processing of the second constituent is easier in the low-frequency condition (and thus, goes faster in the parafovea), one would expect a greater display change effect in this condition than in the high-frequency condition. This hypothesis is of a piece with the well-known finding that predictable words are more likely to be skipped than unpredictable words and, when fixated, have shorter fixation durations (Balota, Pollatsek, & Rayner, 1985; Ehrlich & Rayner, 1981; Rayner & Well, 1996).

We conducted two sets of statistical analyses to test the constraint hypothesis given above. We first examined whether compound family size (here, the number of compounds beginning with the first constituent) may appear as a significant predictor of the relative size of the display change effect (see de Jong, Feldman, Schreuder, Pastizzo, & Baayen, 2002, for the effect of family size on compound processing). As was suggested above, the mean family size was much greater for the high-frequency than for the low-frequency first-constituent words (323 vs. 31). More interesting, this variable predicted the size of the display change condition for the total set of stimuli ($p < .05$). The nature of the effect is such that the smaller the family size, the bigger the change effect. In the analysis of all items, we could not tease apart the effect of family size and that of first-constituent frequency, because these two predictors correlated highly ($r = .83$) with each other. On the other hand, when the high- and the low-frequency words were analyzed separately, the correlations between family size and first-constituent frequency were much lower ($r = .29$ and $.20$, for the low- and high-frequency conditions, respectively). Now, family size was a marginally significant predictor of the size of the display change effect for the low-frequency stimuli ($p = .075$), even when first-constituent frequency was entered in the analysis. The nature of the effect was analogous to that observed for the complete data set: the smaller the family size, the bigger the change effect. For the high-frequency stimuli, family size did not predict the size of the display change effect. This pattern could be explained by positing that family size matters only when it is small (i.e., there is an upper threshold above which family size no longer affects processing; see Moscoso del Prado Martín, Bertram, Häikiö, Schreuder, & Baayen, in press).

Thus, the most promising account of the interaction appears to be the constraint hypothesis: Processing of the second constituents of low-frequency words is easier because they are more constrained and, thus, their processing is interfered with more by the removal of the parafoveal information. The family size measure, which is a plausible index of constraint, appears to predict the

pattern of the effect quite well. However, there is still one problem concerning the constraint hypothesis: The subgaze₂ for low-frequency first constituents was longer than that for high-frequency first constituents, even in the no-change condition. This seems inconsistent with the hypothesis that constraint makes the second constituent easier to process for the low-frequency condition. This led us to conduct two new analyses of subgaze₂.

One possible way to explain this apparent inconsistency is to posit that some late process other than encoding the second constituent—perhaps, rechecking the first constituent or integrating the two components—is more time consuming for the low-frequency first-constituent words than for the high-frequency first-constituent words and causes the increase in subgaze₂. To check out whether this was the case, we divided subgaze₂ into two components: the sum of the fixation durations on the second constituent and the sum of the fixation durations on the first constituent. In fact, the sum of the fixation durations on the second constituent was 18 msec larger for the high-frequency first-constituent words [329 vs. 311 msec; $F_1(1,22) = 10.48$, $MS_e = 695$, $p < .01$; $F_2(1,47) = 2.29$, $MS_e = 4,920$, $p = .14$], whereas the average *revisit* time (no revisit was coded as 0) on the first constituent was larger for the low-frequency first-constituent words [85 vs. 47 msec; $F_1(1,22) = 19.39$, $MS_e = 1,689$, $p < .001$; $F_2(1,47) = 11.62$, $MS_e = 6,149$, $p = .001$]. The latter result is consistent with our prior analysis that showed that people went back to the first constituent more often once having fixated the second constituent, when the first constituent was of low frequency (see Table 2). The results above are also compatible with those reported by Hyönä and Pollatsek (2000). These analyses led us to conclude that the constraint hypothesis is still tenable. In other words, at least a part of the greater display change effect observed in gaze duration for the low-frequency condition is due to the second constituent's being more predictable there than in the high-frequency condition.

DISCUSSION

The present study is, to our knowledge, the first in which eye-movement-contingent display changes were made within words. In prior reading studies, the method was applied to words adjacent to the fixated word. Thus, it is of interest to compare the size of the change effects observed in the present study with those reported earlier. When a display change is introduced in the word to the right of the fixated word, the change effect in gaze duration is typically on the order of 30–40 msec (see Table 3; for a review, see Rayner, 1998; Rayner & Pollatsek, 1989). For example, Rayner, Balota, and Pollatsek (1986) obtained a 40-msec effect for a semantically related but visually dissimilar preview (*song* was initially replaced by *tune*). When the preview was a visually similar nonword (*song* was replaced by *sorp*), the change effect was reduced to a negligible one (5 msec). Our 101-msec overall change effect in gaze duration thus seems to be apprecia-

Table 3
Size of the Parafoveal Preview Benefit (in Milliseconds; Identical – Preview Type) in Previous
Eye-Movement–Contingent Display Change Studies

Study	Type of Parafoveal Preview					
	All Xs	Visually Dissimilar Letters	Visually Similar Letters	First 2–3 Identical + Rest Xs	First 2–3 Identical + Rest Visually Dissimilar	First 2–3 Identical + Rest Visually Similar
Balota, Pollatsek, & Rayner (1985)		+31†	+8†			
Rayner, Balota, & Pollatsek (1986)		+40‡	+5†			
Lima (1987), Experiment 1	+18			+1		
Lima (1987), Experiment 2		+29†			+30‡	
Inhoff (1989a), Experiment 1	+52			+26		
Inhoff (1989a), Experiment 2	+53			+38		
Inhoff (1989b), Experiment 1*	+54			+38		
Inhoff (1989b), Experiment 2*	+32			+24		
Inhoff (1989b), Experiment 3*		+22†			+16†	
Henderson & Ferreira (1990), Experiment 1					+5†	–6†
Henderson & Ferreira (1990), Experiment 2					+11†	+1†
Pollatsek, Lesch, Morris, & Rayner (1992)		+43‡	+25‡		+10‡	–17‡
Briehl & Inhoff (1995), Experiment 1				+38		
Briehl & Inhoff (1995), Experiment 2	+41			+30		
Kennison & Clifton (1995)		+28†				
Inhoff, Starr, & Shindler (2000)		+91†				
Altarriba, Kambe, Pollatsek, & Rayner (2001)		+33§	+15§			
Mean	+42	+41	+15	+28	+14	–7

*Only first fixation durations were reported. †Nonword. ‡Real word. §Not a word in the language of the test sentence.

bly larger than the typical effect sizes observed previously for parafoveal words, especially since the first two letters of the second constituent were preserved in our display change conditions. Thus, within-word letter changes, even if they appear at the end of long compounds and the changes preserve the overall visual shapes of the changed letters, exert large effects on word processing.²

Considering that the effect of display change on the gaze duration was about 100 msec, it is quite remarkable that the effect did not surface when the first part of the compound word was initially fixated, but showed up only later (i.e., when the second constituent was now changed to its correct form). This finding is in line with the serial model of compound word processing we sketched in the introduction. That is, we observed a clear effect for the frequency of the first constituent in early processing measures, but a nonword appearing as the second constituent produced no discernible effects on the following measures: the first-fixation duration, the subgaze for first constituent, or the length of the first within-word saccade. The only indication for a relatively early effect of display change was on the probability of making a second fixation on the first constituent, where a change effect was observed for low-frequency first constituents.

The finding that the change had little influence on the initial stages of compound word processing is also consistent with the serial models of eye guidance in reading, such as the E-Z reader model (Reichle et al., 2003). The basic assumption of the E-Z reader model is that words are processed serially. When enough information from the fixated word is processed so that full lexical access will soon be achieved, a program is made for the next word in the text. An attention shift follows when lexical

access is more complete. It is during this stage (when the eyes are still on word N and attention has shifted to word $N + 1$) that parafoveal preprocessing of $N + 1$ is assumed to take place (for the nature of parafoveal processing, see below). What is crucial in the present context is that parafoveal preprocessing is assumed to have no influence on the processing time devoted to the foveal word; instead, the benefit of parafoveally processing word $N + 1$ is cashed in when it is subsequently fixated. If these assumptions are applied to the processing of long compounds (by assuming that compound word constituents are processed serially), the prediction would be that lexical changes made to the second constituent would not influence processing when the first constituent is fixated but would surface only when the second constituent is fixated. Our results are in line with this prediction.

Secondarily, the present study was also designed to test the notion that processing difficulty affects perceptual span. According to this view, when foveal load is high (as is presumably the case when a low-frequency constituent is fixated), the perceptual span would be constricted, and less parafoveal processing would be done on the second constituent. If this view were correct, we should have obtained a smaller display change effect in the low-frequency condition than in the high-frequency condition. However, exactly the opposite pattern was observed for gaze duration. That is, obscuring the lexical identity of the second constituent early in processing caused an extra delay in processing of the word for low-frequency first-constituent words. Above, we discussed three alternative explanations for the unexpected results, and the constraint hypothesis appeared to be the best candidate. That is, because the second constituents were

more strongly constrained by the first constituent for the low-frequency first-constituent compound words, the second constituent would be easier to process parafoveally in the low-frequency condition. This, in turn, would result in a greater disruption effect when the second constituent is initially replaced by “garbage letters.” The two sets of follow-up analyses provided evidence consistent with this view.

What type of information is picked up from the second constituent while the first constituent is fixated on? Is it visual, orthographic, or lexical? Although we cannot answer this question conclusively, we can rule out the possibility that it will be purely visual in nature. If low-level visual information is what is primarily extracted from the second constituent before it is fixated, we should have obtained only a negligible effect of display change, because the overall visual shape of the changed letters was kept intact in the display change condition (ascenders were replaced with other ascenders, and descenders with other descenders). Thus, by the logic of exclusion, we argue that either orthographic or lexical features related to the second constituent were picked up. Previous research in which parafoveal preprocessing across adjacent words was examined has convincingly shown that both orthographic information (presumably in the form of abstract letter codes) and phonological information were picked up from the parafoveal word, which facilitated the word’s subsequent foveal processing (Rayner, 1998). On the other hand, the evidence in support of lexical-semantic information being picked up is clearly less conclusive (for a recent review of the controversy, see Rayner et al., 2003). Of course, the largely negative evidence observed previously for preprocessing across words does not rule out the possibility that lexical information may, in principle, be picked up from second constituents of compounds. However, at present, we consider it more probable that the useful information is primarily orthographic in nature.

The processing of long two-noun compound words appears to be quite similar to the process of reading two adjacent words separated by a space. That is, our findings indicate that a long compound word is identified by the access of the first constituent before the access of the second constituent is initiated—analogueous to accessing word N , followed by accessing word $N+1$. Moreover, the finding that the ease of processing the second constituent was modulated by the constraint on the second constituent seems to be analogueous to the finding that the predictability of a word from the prior context influences the speed of processing a word. However, whether these two phenomena are really the same is an open question. That is, constraint, as we defined it, is not the same as predictability but is obviously correlated with it. Because it would clearly be very difficult, if not impossible, to come up with objective constraint measures for words in sentences, comparing the two effects will be fairly difficult. However, we think that the size of the display change effect we observed in the present experiment suggests that the constraint of the first constituent on the second constituent in a compound word

is a somewhat different phenomenon than the predictability of a word from sentence context.

As was indicated earlier, the effect of the display change seemed to be much larger in the present experiment than in the typical experiment, where there have been display changes involving entire words. This is documented in Table 3, which is a summary of the size of the parafoveal preview benefit observed in prior eye-movement–contingent display change studies. One might argue that our larger display change effect can be explained by *low-level* factors, such as the distance of the fixation prior to the display change from the changed area or the number of letters changed. We think this is unlikely, since the last fixation on the target word prior to the display change was typically 4–5 characters from the changed letters, which is comparable to previous studies. Similarly, our average display change was about four letters—a figure also quite comparable to that in prior studies in which the whole word was changed. Instead, we think the larger display change effect in the present study is due to “deeper” factors. Note that there is one study in Table 3 (Inhoff, Starr, & Shindler, 2000) that reported an effect of display change comparable to ours (91 msec). An examination of their appendix indicates that almost all their changed words were the second halves of spaced compound words, such as *garage door*. Thus, our large display change effect size is probably not due to our display change being *within visual word* (i.e., the incorrect letters being part of a string of letters that is fixated). Instead, the crucial factor seems to be that our display change is within word in the linguistic sense of *word*. It thus appears that the second constituent of a compound (whether physically joined to the first constituent or not) captures attention earlier and, perhaps, more extensively than does a parafoveal word that stands by itself. This is likely related to the constraint on the second constituent given the first—whether the constituents are physically joined or not. Since the display change effects in the present study and in Inhoff et al. are much larger than those in the conditions in Balota et al.’s (1985) study, where the target word was quite predictable, it appears that within-word constraint is not likely to be explained merely in terms of predictability. In sum, the present study shows that processing constituents in long compound words is, in some sense, serial. However, the fact that the constituents belong to the same word makes them very dependent on each other, and their processing appears to be different from that of successive words.

REFERENCES

- ALTARRIBA, J., KAMBE, G., POLLATSEK, A., & RAYNER, K. (2001). Semantic codes are not used in integrating information across eye fixations in reading: Evidence from fluent Spanish–English bilinguals. *Perception & Psychophysics*, *63*, 875–890.
- ANDREWS, S., MILLER, B., & RAYNER, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology*, *16*, 285–311.
- BALOTA, D. A., POLLATSEK, A., & RAYNER, K. (1985). The interaction

- of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, **17**, 364-390.
- BERTRAM, R., & HYÖNÄ, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from reading short and long Finnish compounds. *Journal of Memory & Language*, **48**, 615-634.
- BRIIHL, D., & INHOFF, A. W. (1995). Integrating information across fixations during reading: The use of orthographic bodies and exterior letters. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 55-67.
- DE JONG, N. H., FELDMAN, L. B., SCHREUDER, R., PASTIZZO, M., & BAAYEN, R. H. (2002). The processing and representation of Dutch and English compounds: Peripheral morphological and central orthographic effects. *Brain & Language*, **81**, 555-567.
- DEUTSCH, A., FROST, R., PELLE, S., POLLATSEK, A., & RAYNER, K. (2003). Early morphological effects in reading: Evidence from parafoveal preview benefit in Hebrew. *Psychonomic Bulletin & Review*, **10**, 415-422.
- EHRICH, S. F., & RAYNER, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning & Verbal Behavior*, **20**, 641-655.
- 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, & Cognition*, **16**, 417-429.
- HYÖNÄ, J. (2003). *Morphological preprocessing of first constituents of long compounds*. Manuscript in preparation.
- 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). Morphological 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). Oxford: Elsevier.
- INHOFF, A. W. (1989a). Lexical access during eye fixations in reading: Are word access codes used to integrate lexical information across interword fixations? *Journal of Memory & Language*, **28**, 444-461.
- INHOFF, A. W. (1989b). Parafoveal processing of words and saccade computation during eye fixations in reading. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 544-555.
- INHOFF, A. W., BRIIHL, D., & SCHWARTZ, J. (1996). Compound word effects differ in reading, on-line naming, and delayed naming tasks. *Memory & Cognition*, **24**, 466-476.
- INHOFF, A. W., RADACH, R., & HELLER, D. (2000). Complex compounds in German: Interword spaces facilitate segmentation but hinder assignment of meaning. *Journal of Memory & Language*, **42**, 23-50.
- INHOFF, A. W., STARR, M., & SHINDLER, K. L. (2000). Is the processing of words during eye fixations in reading strictly serial? *Perception & Psychophysics*, **62**, 1474-1484.
- JUHASZ, B. J., STARR, M. S., INHOFF, A. W., & PLACKE, L. (2003). The effects of morphology on the processing of compound words: Evidence from naming, lexical decisions and eye fixations. *British Journal of Psychology*, **94**, 223-244.
- KENNEDY, A., PYNTE, J., & DUCROT, S. (2002). Parafoveal-on-foveal interactions in word recognition. *Quarterly Journal of Experimental Psychology*, **55A**, 1307-1337.
- KENNISON, S. M., & CLIFTON, C. (1995). Determinants of parafoveal preview benefit in high and low working memory capacity readers: Implications for eye movement control. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 68-81.
- LAINÉ, M., & VIRTANEN, P. (1999). *WordMill Lexical Search Program*. Turku, Finland: University of Turku, Center for Cognitive Neuroscience.
- LIMA, S. D. (1987). Morphological analysis in sentence reading. *Journal of Memory & Language*, **26**, 84-99.
- MOSCOSO DEL PRADO MARTÍN, F., BERTRAM, R., HÄIKIÖ, T., SCHREUDER, R., & BAAYEN, R. H. (in press). Morphological family size in a morphologically rich language: The case of Finnish compared to Dutch and Hebrew. *Journal of Experimental Psychology: Learning, Memory, & Cognition*.
- POLLATSEK, A., & HYÖNÄ, J. (in press). The role of semantic transparency in the processing of Finnish compound words. *Language & Cognitive Processes*.
- POLLATSEK, A., HYÖNÄ, J., & BERTRAM, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human Perception & Performance*, **26**, 820-833.
- POLLATSEK, A., LESCH, M., MORRIS, R. K., & RAYNER, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 148-162.
- RAYNER, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, **7**, 65-81.
- RAYNER, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, **124**, 372-422.
- RAYNER, K., BALOTA, D. A., & POLLATSEK, A. (1986). Against parafoveal semantic preprocessing during eye fixations in reading. *Canadian Journal of Psychology*, **40**, 473-483.
- RAYNER, K., & POLLATSEK, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice-Hall.
- RAYNER, K., & WELL, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, **3**, 504-509.
- RAYNER, K., WHITE, S. J., KAMBE, G., MILLER, B., & LIVERSEGE, S. P. (2003). On the processing of meaning from parafoveal vision during eye fixations in reading. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 213-234). Amsterdam: Elsevier Science.
- 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.
- REICHLE, E. D., RAYNER, K., & POLLATSEK, A. (2003). The E-Z reader model of eye movement control in reading: Comparisons to other models. *Behavioral & Brain Sciences*, **26**, 445-526.

NOTES

1. In order to treat the no-change trials comparably, the same exclusion procedure was also applied to them. In the no-change condition, an identical frame was redrawn during the critical saccade (it should be noted that the redrawing cannot be perceived in any circumstances). Of the trials, 20.2% were excluded.

2. One may argue that the sizable display change effect observed in the present study for a within-word change was due to more letters being changed than in the previous studies. However, this may not be the case, since Hyönä (2003) found only a 13-msec cost on gaze duration on 12- to 14-letter Finnish compound words when all but the first 3-4 letters were replaced by random letters before the words were fixated.

(Manuscript received May 16, 2003;
revision accepted for publication November 12, 2003.)