LANGUAGE AND COGNITIVE PROCESSES
2005, 20 (1/2), 261–290

Psychology Press
Taylor & Francis Group

The role of semantic transparency in the processing of Finnish compound words

Alexander Pollatsek
University of Massachusetts, Amherst, MA, USA

Jukka Hyönä
University of Turku, Finland

Three experiments examined whether the semantic transparency of a long Finnish compound word has any influence on how the compound word is encoded in reading. The frequency of the first constituent (as a separate word) was manipulated, while matching for the frequencies of the compound word and of the second constituent. The effect of this frequency manipulation on encoding time served as a 'marker' that the compound word was processed, at least in part, componentially. In Experiment 1, each high-frequency transparent compound was paired with a low-frequency transparent compound, and each high-frequency opaque compound was paired with a low-frequency opaque compound. A sentence frame was created for each pair that was identical up to the word following the target word. In Experiments 2 and 3, the matching was done between transparent and opaque word pairs. In addition, Experiment 3 had a display change manipulation in which most of the second constituent was not visible until it was fixated. Readers' eye fixation patterns on and immediately after the target word were examined. Reliable first constituent frequency effects were observed in the fixation duration measures on the target word, but there were no effects of transparency. In addition, a comparison of the display change condition to the standard condition indicated that the constituents of the compound word were processed sequentially. It thus appears that the

Correspondence should be addressed to: Alexander Pollatsek, Department of Psychology, University of Massachusetts, Amherst MA 01003, USA.
Email: pollatsek@psych.umass.edu.

The first author was supported by Grants HD17246 and HD26765 from the National Institutes of Health and the second author acknowledges the support of Suomen Akatemia (the Academy of Finland). We thank Jonathan Grainger, Gary Libben, and an anonymous reviewer for their comments on a previous version of this article. We are also grateful to Tuomo Häikiö for his help in analysing the data of Experiment 3.

http://www.tandf.co.uk/journals/pp/01690965.html DOI: 10.1080/01690960444000098

identification of both transparent and opaque long compound words takes place, at least in part, by accessing the constituent lexemes and does not rely on constructing the meaning from the components.

The goal of the present study was to investigate the effect that the semantic transparency of a compound word has on processing a compound word when people are reading sentences. A compound word is usually defined as transparent when the meaning of the compound word is consistent with the meanings of the constituents (e.g., *carwash*). In contrast, a compound word is defined as semantically opaque, when its meaning cannot be constructed by directly combining the meanings of the individual constituents (e.g., *pineapple*).[1]

Our previous work (Bertram & Hyönä, 2003; Hyönä & Pollatsek, 1998, 2000; Pollatsek, Hyönä, & Bertram, 2000), examined the processing of semantically transparent compound words when they are embedded in a sentential context. The participants read the sentences silently for comprehension while their eye fixation patterns were recorded. This technique allowed us to obtain a detailed picture of the time-course of compound word processing during sentence processing. One of the key questions in this research was the extent to which compound words are recognised via their constituents. One way we studied this question was by manipulating the frequency of compound word constituents (i.e., their frequencies as separate words) while matching the compound words on their (whole-word) frequency. For the two-constituent compound words we studied, we observed large and reliable effects of the frequency of both constituents in the eye fixation patterns. The frequency of the first constituent exerted a large effect on the gaze duration on the word (i.e., the sum of the durations of the fixations on the word prior to the eyes leaving the word) and part of this effect was 'early', as the duration of the initial fixation on the compound word was influenced by this manipulation. The frequency of the second constituent also had a large and reliable effect on the gaze duration, but its effect did not appear until the second fixation on the word. These results, which have recently been replicated and extended to English (Andrews, Miller, & Rayner, 2004; Juhasz, Starr, Inhoff, & Placke, 2003), clearly show that compound words are identified, at least in part, via their constituents.

[1] Usually, for most transparent compound words, the meaning of the word cannot be uniquely computed from the constituents, as a *carwash* could be some sort of device that washes with a car; instead the meaning is usually a highly plausible combination of the constituent meanings.

The ‘whole-word’ frequency of the compound, however, also exerts a large and reliable effect on compound word processing. Pollatsek et al. (2000) observed a whole-word frequency effect that emerged relatively early in the eye movement records, and there was even a suggestion of an effect in the first fixation duration. On the basis of these results, Pollatsek et al. proposed a parallel dual-route model for processing compound words in text in which compounds are identified by simultaneously accessing the constituent lexemes and the whole-word representation. The access of the first constituent was assumed to occur prior to accessing either the second constituent or the whole-word representation. More recently, Bertram and Hyönä (2003) demonstrated that the length of a compound word determines which of the two routes (decomposition vs. whole word) is more influential in processing. They showed that with long compounds (about 13 letters), as those used in our previous work, the initial constituent gets a head-start, whereas with shorter compounds (7–9 letters) the whole-word representation receives an early activation. In Experiment 1, there was an early effect of first constituent frequency (as reflected in the first fixation duration) for long compounds, but not for short compounds. In contrast, when whole-word frequency was manipulated (Experiment 2), early processing stages were reliably affected for short compounds but not for long compounds. Bertram and Hyönä felt that their pattern of results was most parsimoniously explained by visual acuity constraints. That is, when a long compound word is initially fixated (usually not far from the middle of the initial constituent), the initial constituent is much easier to process than the entire compound, whereas with shorter compounds the entire word is readily perceivable during the initial eye fixation and thus the whole-word route is speeded relative to the compositional route.[2]

In the present study, we addressed the question of whether long compound words are accessed via their constituents in silent reading even when the meaning of the compound word is not readily constructed as a combination of the constituent meanings. To our knowledge, there are no previous reading studies that address this question. On the other hand, there is a lot of prior work on the role of transparency in compound word processing using the lexical decision paradigm. This research has examined whether transparency modulates the ability of components of compound words either to prime or to be primed using both constituent priming and

---

[2] Some recent attempts to model our data quantitatively (Pollatsek, Rayner, & Reichle, 2003) suggest that the two routes are not likely to be independent, as in a standard ‘race’ model. That is, the sizes of the frequency effects observed in the Finnish compound word experiments were virtually impossible to fit with any reasonable assumptions about processing times for the words or the constituents.

semantic priming techniques (for the effects of semantic opacity in processing inflected words, see Schreuder, Burani, & Baayen, 2003). We will first review the constituent priming studies and then the semantic priming studies.

In the constituent priming studies, a compound word constituent has either been employed as the prime or the target (the studies cited below used the immediate unmasked priming paradigm, except for Monsell (1985) who used long lag priming). Zwitserlood (1994, Experiment 1) used the compound word as the prime and either the first or second constituent as the target, and observed a constituent priming effect of similar magnitude for transparent and opaque Dutch compounds. Monsell (1985) and subsequently Libben and colleagues (Libben, Gibson, Yoon, & Sandra, 1997, 2003; Jarema, Busson, Nikolova, Tsapkini, & Libben, 1999) used one of the compound word constituents as the prime for the compound word and obtained similar results for English compounds. Monsell (1985) observed a healthy priming effect for both transparent and opaque compounds, and Libben et al. (1997, 2003) also observed equal priming effects for opaque and transparent compounds. Libben et al. also distinguished between *fully opaque* compounds such as *humbug* (where the meaning of neither constituent is related to the meaning of the compound) and *partially opaque* compounds such as *strawberry* (where the meaning of one of the constituents is related to the meaning of the compound) and found no difference in priming. Jarema et al. (1999) employed the same constituent priming paradigm to study the recognition of French compounds and found a strong priming effect of similar magnitude for transparent and opaque compounds. In a second experiment employing Bulgarian compounds, they observed a reliable priming effect for all compound word types except for the fully opaque ones.

In sum, constituent priming of similar magnitude was observed for transparent and opaque compounds with one exception: Jarema et al. (1999) did not observe constituent priming for fully opaque Bulgarian compounds. With this possible exception, the constituent priming studies thus indicate that, at some level, the lexical entries of constituents are activated for opaque compounds and are not suppressed. However, it is possible that a different pattern of results might emerge with a task such as semantic priming, which may tap the activation of the meanings of the constituents and not merely their lexical activation. To our knowledge, only two semantic priming studies have examined the effect of semantic transparency on compound word processing. Sandra (1990) primed a compound word (e.g., *milkbottle*) with a lexical item that was semantically associated with the compound word constituent (e.g., *cow*). He observed a reliable priming effect in Dutch for transparent compounds (e.g., *koe* → *melkfles* = cow → milkbottle) but not for opaque compounds (e.g., *koe* →

melkweg = cow → milky way). Zwitserlood (1994) used transparent and opaque Dutch compounds as primes and semantic associates of either the first or second constituent as targets (e.g., transparent: *kerkorgel* → *priester* = church organ → priest; opaque: *drankorgel* → *bier* = drunkard → beer). She found a semantic priming effect for transparent and partly opaque compounds but not for completely opaque compounds. Thus, the two semantic priming studies offer suggestive evidence that the meaning of constituents of opaque compound words may not be available during compound word processing, either because they are not activated or that they are quickly suppressed by the whole-word meaning.

Although the overall pattern of results emerging from the priming studies reviewed above is not totally consistent, one may tentatively conclude that compound word constituents are represented at the lexical level regardless of the semantic transparency of the compound, whereas constituent meanings may not be available for (fully) opaque compounds. In the model proposed by Libben (1998), the mental organisation of transparent and opaque compounds is worked out along these lines. The model distinguishes three levels of representation: stimulus, lexical, and conceptual. At the conceptual level, semantic transparency determines the extent to which constituent meanings are mentally represented. For example, at this level, *bird* in *jailbird* is not mentally represented, whereas *jail* in *jailbird* is. Although Libben's (1998) model posits a qualitative difference in the mental organisation of transparent and opaque compounds at the conceptual level, transparency is not assumed to matter at the lexical level. Both transparent and opaque compounds are assumed to be decomposed into their morphological constituents. Thus, at the lexical level both the constituents and the full-form are represented.

There are two key related questions that we hoped to address with the current research. The first is whether opaqueness produces a processing cost in identifying long Finnish compound words in reading. Although most of the above research suggests that transparency of a compound word did not have a significant role in its priming characteristics, we think it is at least reasonable to suspect that transparency may play a larger role in Finnish. That is because compounding is quite productive in Finnish, and furthermore, as in German, compounds must be written without spaces between the constituents. As a result, it is likely that the typical Finnish reader encounters 10–20 novel compound words a day. Moreover, logically, such novel compounds need to be comprehended by composing the meaning of the constituents, as there is no lexical entry for the word. Because readers cannot know, a priori, whether a letter string is a novel compound or not, it would seem reasonable that at least part of the process of encoding an existing compound would be encoding the meaning of the constituents, and to the degree to which the combination of these meanings differ from the meaning of the compounds, there should be some cost in encoding the word.

The related question is whether the frequency effects we had previously obtained with semantically transparent words would extend to semantically opaque words. At one extreme, if the compositional process in reading is completely at the conceptual level, one might expect the frequency of the initial constituent to be irrelevant for opaque words, as they would have to be recognised by a non-compositional process. Indeed, it is even possible, given this kind of model, that there could be a reverse frequency effect (i.e., longer recognition times when the first constituent is higher), especially if the first constituent is the opaque constituent. That is, it is possible that the higher the frequency of the first (opaque) constituent, the more it will interfere with the correct computation of the compound word's meaning. At the other extreme, if the compositional process is completely at the lexical level, then one might expect the transparency of the compound to have no effect on processing, and thus observe equal frequency effects for transparent and opaque compounds. Another possibility, of course, is that the compositional process is at both levels. If so, one might expect some initial constituent frequency effect for opaque compounds, but less of an effect than for transparent compounds.

In the present study, we examined the time course of compound word processing by examining readers' eye movements as they read both transparent and opaque compound words in sentential context. In three experiments, pairs of compound words, either opaque or transparent, were selected that differed in the frequency of the first constituent but were matched on their second-constituent frequency and whole-word frequency. A sentence frame was constructed for a pair of matched target words that was identical up to the word following the target word. In Experiment 1, either (a) a transparent word with a low-frequency first constituent was paired with a transparent word with a high-frequency first constituent or (b) an opaque word with a low-frequency first constituent was paired with an opaque word with a high-frequency first constituent. This was to assess whether both classes of words produced first-constituent frequency effects, and if so, whether they were of differing magnitudes. In Experiments 2 and 3, an opaque compound was directly matched with a transparent one of the same first constituent frequency class to provide a more sensitive test of whether transparency had any effect on reading. Participants read the sentences for comprehension, while their eye fixation patterns were registered. (We will describe other details of Experiment 3 later.)

Eye-movement measures provide the researcher with a detailed picture of the time-course of processing (for a review, see Rayner, 1998). This should be helpful in sorting out lexical and semantic effects, if one assumes that the lexical effects occur earlier. That is, if we find that both lexical and compositional effects occur, one might expect to find the effects due to the frequency of the first constituent early in processing, but not later in processing. We used as early processing measures the duration of initial fixation on the target word and the probability of making a second fixation on the word (the decision to refixate the word needs to be made during the initial fixation). As somewhat later processing measures, we employed the gaze duration on the target word, the durations of the second and third fixations, and the probability of making three fixations on the word. As measures of even later processing, we employed the probability of making a regression out of the target word and the probability of skipping word $N + 1$.

EXPERIMENT 1

Method

Participants. Twenty-six students from the introductory psychology course at the University of Turku, all of whom were native speakers of Finnish, participated in the experiment as a part of the course requirement.

Apparatus. Eye movements were collected by the EYELINK eye-tracker manufactured by SR Research Ltd. (Canada). The eyetracker is an infra-red video-based tracking system combined with hyperacuity image processing. There are two cameras mounted on a headband (one for each eye) including two infra-red LEDs for illuminating each eye. (The headband weighs 450 g.) The cameras sample pupil location and pupil size at the rate of 250 Hz. Registration is binocular and is performed for the selected eye(s) by placing the camera(s) and the two infra-red light sources 4–6 cm away from the eye. In the present study, registration was done monocularly using the right eye. The resolution of eye position is 15 seconds of arc and the average spatial accuracy approximately 0.5 degrees. Head position with respect to the computer screen is tracked with the help of a head-tracking camera mounted on the centre of the headband at the level of the forehead. Four LEDs are attached to the corners of the computer screen, which are viewed by the head-tracking camera, once the participant sits directly facing the screen. Possible head motion is detected as movements of the four LEDs and is compensated for on-line from the eye position records.

Materials and design. A set of 40 semantically transparent and 40 opaque compounds was selected. All target words were 2-noun compound words (12–15 characters long). For both sets, the frequency of the first constituent was manipulated: one set had a frequent first constituent and the other set had an infrequent first constituent. There were 20 words of each kind both for transparent and opaque words (i.e., a total of 80 target items). 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 (1996).

By definition, a semantically transparent compound is one whose approximate meaning can be derived from the constituent meanings by 'glueing' them together. On the other hand, a compound word is semantically opaque when its meaning cannot be computed by simply glueing together constituent meanings. In our set of opaque compounds either the meaning of the whole word was opaque (e.g., *kompastuskivi* = stumbling block), or the meaning of the first constituent was opaque (e.g., *verivihollinen* = blood enemy) and the meaning of the second constituent was transparent (none of the words was opaque only in its second constituent). In the former case, the compound word meaning was often metaphorical. An initial selection of the target words was done by intuition. To ensure that the words indeed differed in semantic transparency we asked a group of eight subjects to rate the words for transparency using a 7-point scale (1 = totally transparent, 7 = totally opaque). As the ratings indicated (see Table 1), the two word sets clearly differed in their perceived transparency.

TABLE 1
Characteristics of the target words used in Experiment 1, 2, and 3

	Transparent compounds		*Opaque compounds*	
Stimulus characteristic	*Low frequency first constituent*	*High frequency first constituent*	*Low frequency first constituent*	*High frequency first constituent*
Frequency of first constituent[a,b]	5	274	5.7	299.1
Frequency of compound word[a,c]	2	2.4	2.2	2.2
Frequency of second constituent[a,b]	229.5	189.3	168.9	59.5
Length of compound word (in characters)	12.9	12.8	13.2	12.9
Length of first constituent	7.2	6.1	6.6	5.9
Rated semantic transparency[d]	1.52	1.6	5.24	4.76

[a] These frequencies have been converted to frequencies per million.
[b] The frequencies are of the constituents as separate words in all their inflectional forms, including the nominative case.
[c] The frequencies are of the compound words in all their inflectional forms, including the nominative case.
[d] 1 = completely transparent, 7 = completely opaque.

Characteristics of the target words are presented in Table 1. There was a large difference in the frequency of the first constituent between the two word sets, whereas the whole-word frequency was closely matched. The target words were also matched for word length and for first constituent length.

The target words were near the beginning of sentences, but never in the initial word position. A 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 so that it was identical up through the word following the target word. The matching was done separately for transparent and opaque compounds. A typical sentence pair is as follows, with the target word in italics.

Low-frequency first constituent condition, transparent compound

Edessäni huomasin *alttaritaulun*, joka oli niin täynnä yksityiskohtia, etten jaksanut tutkia sitä sen tarkemmin.
(In front of me I saw an *altarpiece*, which was so full of detail so that I didn't have the energy to examine it any closer.)

High-frequency first constituent condition, transparent compound

Edessäni huomasin *yleisöjoukon*, joka oli varmaan menossa samaan konserttiin.
(In front of me I saw a *spectator group* that was presumably going to the same concert.)

To ensure that the two sentences of each pair were equally natural in meaning, we asked a separate group of eight subjects to compare the naturalness of the two sentences in each pair. The subjects were asked to choose between three alternatives: (1) the sentences are equally natural, (2) sentence A is more natural than sentence B, and (3) sentence B is more natural than sentence A. If the majority of subjects favoured one sentence over the other, that sentence pair was revised, and another group of three subjects performed the comparison for ten revised sentence pairs. The outcome of this procedure was that the final set of sentence pairs were equated for the perceived naturalness.

The target sentences were presented in Courier font starting from a point near the left of the computer screen. The sentences occupied a maximum of three lines of text, and the critical word never appeared as the initial or final word of a text line. With a viewing distance of about 65 cm, one character subtended approximately 0.5 degrees of visual angle in the horizontal direction. There were two blocks of sentences, with the sentence frames containing each word pair appearing in separate blocks. The order of the blocks was counterbalanced across participants. Thus, each participant saw all the critical target words. Within each block, the order of target sentences was randomised. There were 30 filler sentences among the critical sentences.

Procedure. The eye-tracker was first calibrated using a nine-point calibration grid which spanned the entire screen area. Ten practice sentences were presented before the actual experiment. Participants were instructed to read each sentence for comprehension. They were told that they would be occasionally asked to paraphrase the sentence they had just finished reading. Before the presentation of each sentence, they were required to gaze at a fixation point in the top left corner of the screen. In case of a calibration drift, the calibration was automatically corrected. Immediately after the fixation point was erased from the screen, a sentence, extending 1–3 lines of text, was presented left-justified and centred vertically on the screen. The experimental session took a maximum of 30 minutes.

Results and discussion

Gaze duration. The most common measure of global processing time on the target word is *gaze duration*, the sum of the durations of all fixations on the first encounter with the target word (i.e., prior to a saccade leaving the target word). As seen in Table 2, there was a 40 ms difference in gaze durations between the target words with high frequency first constituents and those with low frequency first constituents, $F_1(1, 24) = 28.3$, $p < .001$, $F_2(1, 38) = 6.02$, $p < .05$. Surprisingly, there was no effect of transparency; the overall difference between transparent and opaque compound words was 1 ms, and the interaction between transparency and frequency was very small, Fs < 1.

More detailed measures on the target word. A more detailed account of processing can be obtained by looking at individual fixations (see Table 2). First, consider the duration of the initial fixation on the word. Here (quite surprisingly in view of the above null result for transparency), there was an indication of a main effect of transparency, $F_1(1, 24) = 20.7$, $p < .001$, $F_2(1, 38) = 3.42$, $p < .10$. It should be noted that this effect may be a result of the different sentence frames used for transparent and opaque compounds (cf. Experiments 2 and 3). The small effect of the first constituent frequency was not significant overall, $F_1(1, 24) = 1.50$, $p > .20$, $F_2 < 1$, nor was the 5 ms frequency effect for the transparent compounds, $t_1(24) = 1.89$, $p < .10$, $t_2(24) = 1.05$, $p > .20$. There was a slight hint of a frequency effect in the duration of the second fixation, although all Fs in

TABLE 2
Experiment 1: Various reading indexes on the target word as a function of first constituent frequency and semantic transparency

	Transparent compounds		*Opaque compounds*	
Eye movement measure	*Low frequency first constituent*	*High frequency first constituent*	*Low frequency first constituent*	*High frequency first constituent*
Gaze duration (in milliseconds)	538	492	540	493
First fixation duration (in milliseconds)	193	188	199	200
Second fixation duration (in milliseconds)	191	188	187	187
Third fixation duration (in milliseconds)	182	184	184	177
Probability of an initial fixation	1	0.98	0.98	0.98
Probability of refixating target word	0.93	0.9	0.92	0.88
Probability of refixating target word at least twice	0.51	0.46	0.55	0.45
Per cent of saccades leaving target word that are regressions	3.2	3.2	6	5.8
Per cent of forward saccades that skip word N + 1	3.3	3.4	6.2	6.2

the analyses were less than 1. Thus, little of the frequency effect in the gaze duration is attributable to differences in individual fixation durations. Instead, most of the frequency effect on the gaze durations appears to be due to differing numbers of fixations made on the target word. The probability of making a second fixation was 3.3% greater for the low frequency words, $F_1(1, 24) = 7.73$, $p < .01$, $F_2(1, 38) = 5.59$, $p < .05$, and the probability of making a third fixation was 5.2% greater for the low frequency words, $F_1(1, 24) = 19.1$, $p < .001$, $F_2(1, 38) = 6.45$, $p < .05$. None of the transparency effects on the probability of refixation were significant.

Effects subsequent to processing the target word ('spillover'). As transparency may have 'delayed' effects, we thought it would be of interest to examine where the eyes went after leaving the target word. Unfortunately, these data are not completely 'clean' as the text prior to and after the target word were not the same for the transparent and opaque words. We first examined the per cent of times that people left the target word by regressing back to a prior word. In fact, the number of regressions from opaque words was almost double the number of regressions from transparent words (6.0% vs. 3.2%), $F_1(1, 24) = 8.77$,

$p < .01$, however, $F_2(1, 38) = 3.10$, $p < .10$. A second measure that was taken was the probability that the reader would skip the word after the target word when leaving the target word. Here, surprisingly, there appeared to be a difference in the 'wrong' direction, as there were more such skips for opaque words than for transparent words, $F_1(1, 24) = 8.79$, $p < .01$, however, $F_2(1, 38) = 3.12$, $p < .10$. However, again, these two differences could have been due to differences in the sentence frames.

EXPERIMENT 2

Although transparency clearly had no effect on the total 'first pass' processing time on the target word (as measured by the gaze duration) in Experiment 1, there were some suggestions that transparency might have had some subtle effects (e.g., on the duration of the first fixation). There is a problem with those comparisons, however, because the matched transparent and opaque words were in different sentence frames. Therefore, these transparency effects could merely have been due to uncontrolled differences in the sentence frames preceding the target words. As a result, in Experiment 2, we used matched transparent and opaque compound words in the same sentence frames to get a better measure of transparency effects.[3]

Method

Participants. Twenty-four students from the introductory psychology course at the University of Turku, all of whom were native speakers of Finnish, participated in the experiment as a part of the course requirement. None of them took part in Experiment 1.

Apparatus. The same apparatus was used as in Experiment 1.

Materials and design. The same target words were used as in Experiment 1. However, the target words appeared in different sentence frames than in Experiment 1. The matching of sentence frames was performed by either pairing an opaque and transparent compound that had a high-frequency initial constituent or pairing an opaque and transparent compound that had a low-frequency initial constituent. This allowed us to make more reliable comparisons between opaque and

[3] The main reason we did not use a completely crossed design (which would have allowed transparency and frequency to both be within-item variables) is that the resulting sentences would have been much less natural. That is, it is quite difficult to construct a sentence frame such that four compound words can equally felicitously fit in. Further, in such a design, we would have had to construct four such frames for each quartet of target words.

transparent compounds. Each participant saw 120 sentences: the 80 experimental sentences containing the target words and 40 filler sentences.

A group of five subjects performed naturalness ratings for each sentence pair as described above for Experiment 1.

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

Results and discussion

Gaze duration. As in Experiment 1, we start with *gaze duration*, the sum of the durations of all fixations on the first encounter with the target word (i.e., prior to a saccade leaving the target word). As seen in Table 3, the pattern was quite similar to that in Experiment 1. There was a 36 ms difference in gaze durations between the target words with high frequency first constituents and those with low frequency first constituents, $F_1(1, 23) = 26.5$, $p < .001$, however, $F_2 < 1$. This frequency effect was not reliable across items largely because frequency was a 'between item' manipulation in this experiment in order to allow for the most reliable estimates of

TABLE 3
Experiment 2: Various reading indexes on the target word as a function of first constituent frequency and semantic transparency

	Transparent compounds		*Opaque compounds*	
Eye movement measure	*Low frequency first constituent*	*High frequency first constituent*	*Low frequency first constituent*	*High frequency first constituent*
Gaze duration (in milliseconds)	457	432	479	433
First fixation duration (in milliseconds)	205	203	209	206
Second fixation duration (in milliseconds)	196	181	189	183
Third fixation duration (in milliseconds)	199	177	180	179
Probability of an initial fixation	0.98	0.99	1	0.98
Probability of refixating target word	0.8	0.78	0.8	0.75
Probability of refixating target word at least twice	0.31	0.3	0.35	0.3
Per cent of saccades leaving target word that are regressions	4.5	8	5.6	6.7
Per cent of forward saccades that skip word N + 1	32.1	43.4	33.2	41.5
Duration of first fixation on Word N + 1 (in milliseconds)	196	199	204	205

transparency effects. Nonetheless, in spite of the greater power and control of the current experiment for analysing transparency, there was only a suggestion of a transparency effect (11 ms) which was not close to being significant, $F_1(1, 23) = 2.00$, $p > .10$, $F_2 < 1$. Unlike in Experiment 1, there was a suggestion of an interaction between transparency and frequency, but the effect was small, $F_1(1, 23) = 1.90$, $p > .10$, $F_2 < 1$, and it was in the opposite direction of what one might expect—with a larger first constituent frequency effect for the opaque words.

More detailed measures on the target word. The only significant effect of constituent frequency on individual fixation durations was the 11 ms difference in second fixation duration, $F_1(1, 23) = 9.34$, $p < .01$, $F_2(1, 38) = 4.52$, $p < .05$. None of the main effects of transparency on individual fixation durations was close to significant (all ps $> .20$ and most Fs < 1) and the only effect involving transparency that was close to significant was the interaction suggesting that the frequency effect was larger for transparent words, $F_1(1, 23) = 3.81$, $p < .10$, $F_2(1, 38) = 2.46$, $p > .10$, $F_1(1, 19) = 3.38$, $p < .10$, $F_2 < 1$, for second and third fixation durations, respectively.[4] There were no completely reliable frequency effects on the probability of refixating, possibly due to the fact that this was a 'between-item' variable. The only effect that was close to significant was that the probability of making a first refixation on the target word was 4.2% greater for the low frequency words, $F_1(1, 23) = 4.70$, $p < .05$, but $F_2 = 1$. No effect involving transparency was close to significant (again, all ps $> .20$ and most Fs < 1).

This general pattern thus differs somewhat from Experiment 1. In Experiment 1, little of the difference in gaze duration between high- and low-frequency first constituent words was attributable to differences in individual fixation durations and most of the difference was attributable to differences in refixation probabilities. In contrast, in Experiment 2, roughly half of the gaze duration difference appeared to be due to differences in the second and third fixation durations.

Measures after leaving the target word ('spillover'). The sentence frames containing the pairs of matched transparent and opaque words were identical up to the word following the target word. Thus, it is of interest to see whether transparency did have some effect immediately

[4] It is highly unlikely that the trend for the frequency effect being slightly larger in the duration of second and third fixation for transparent compounds is a result of the second constituent in the low-frequency opaque condition being of lower frequency than in the other conditions. This is because the trend is primarily due to relatively long fixations observed in the low-frequency transparent condition.

after leaving the target word. Perhaps the purest measure is the per cent of time that people initially left the target word with a regression back to prior text. Here, somewhat surprisingly, there was a suggestion of a frequency effect, $F_1(1, 23) = 3.94$, $p < .10$, but $F_2 < 1$, which was in the 'wrong' direction, with more regressions when the initial constituent was high frequency. However, there was no effect of transparency, with *F*s close to zero. A second measure of 'where' the eyes landed after leaving the target word is the per cent of time people skipped word N + 1 and went immediately to word N + 2. In fact, people skipped word N + 1 almost 10% more in the high frequency condition, but this effect was not reliable over items, $F_1(1, 23) = 29.5$, $p < .001$, but $F_2(1, 38) = 1.20$, $p > .20$. Again, there was no effect of transparency, *F*s < 1. A third measure to assess processing immediately after leaving the target word was the duration of the first fixation after leaving the target word for those fixations that landed on word N+1. Here was one of the few indications of a transparency effect, as the mean fixation duration was 7 ms less for transparent than for opaque words, $F_1(1, 23) = 3.79$, $p < .10$, but $F_2 < 1$.

Summary. Experiment 2 replicated the main features of Experiment 1. Again, there was no significant transparency effect on gaze duration, and the few somewhat odd transparency effects that occurred on detailed measures in Experiment 1 disappeared, indicating that they were likely due to the differences in sentence frames in Experiment 1 between the opaque and transparent compound words. We observed a first constituent frequency effect in Experiment 2 that was only a little smaller than the one in Experiment 1 (36 ms vs. 45 ms), but it was not significant over items, most likely because constituent frequency was a between-item manipulation in Experiment 2. Finally, there was a marginal tendency (clearly non-significant in the item analysis) observed in the durations of the second and third fixation for the frequency effect to be smaller for the opaque compounds than for the transparent compounds. However, as will become apparent, this tendency was not confirmed in Experiment 3.

EXPERIMENT 3

In Experiment 3, an eye movement contingent display change technique was applied in yet another attempt to uncover effects of semantic transparency on compound word processing. In the experiment, there were two display change conditions. In one, all but the first two letters of the second constituent were replaced with visually similar letters before the second constituent was fixated (which made the second constituent initially a non-word), and in the other (a control condition), there were no display changes. An invisible 'boundary' was set at the constituent boundary; when the eyes crossed this boundary, the changed letters were replaced with the correct ones during the saccadic eye movement. As vision is reduced during saccades (so-called saccadic suppression), readers are not able to see the actual change taking place (provided that the change is completed before the end of the critical saccade). Thus, when the readers fixated on the second constituent, it always appeared in its correct form, and, as indicated above, they were unaware of the display change.

This boundary technique, which was employed previously with transparent compound words (Hyönä, Bertram, & Pollatsek, in press), allowed us to test whether the processing of opaque compounds would be more strongly affected by the display change than the processing of transparent compounds. In other words, if the identification of opaque compounds is less compositional than for transparent compounds and thus is more through the whole-word form, the processing of opaque compounds should be affected more by display change. That is, in the display change condition, the second constituent is initially obscured and the readers can only have access to the first constituent. As a result, the reader would be induced to adopt a compositional processing strategy, which should be less suitable for opaque compounds, if the reader prefers to access opaque compounds as wholes. Furthermore, as most of the opaque compounds in the present study had first constituents that were semantically opaque, such compositional processing may lead to 'garden-path' effects (if the meanings of constituents are being composed).

A second reason for conducting Experiment 3 was to see whether the suggestive trends involving transparency that were observed in gaze duration and in the duration of second and third fixations in Experiment 2 would obtain stronger empirical support. Experiment 3 employed the same materials as Experiment 2 (i.e., the pairwise matching was done between transparent and opaque compounds) with half of the target compounds appearing in the display change condition and the other half in the no change condition.

Method

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

Apparatus. The apparatus used in Experiment 1 and 2 was updated to a second generation EyeLink tracker (EyeLink II, Toronto, Ontario, Canada), which runs at a sampling rate of 500 Hz (i.e., a new sample is recorded every 2 ms).

Materials and design. The same target words were used as in Experiment 1 and 2. The target sentences were the ones used in

Experiment 2, which means that each transparent compound was matched pairwise with an opaque compound. The target compound appeared initially either unchanged (i.e., the second constituent was present throughout the sentence presentation) or the last letters of the second constituent were replaced with visually similar letters (i.e., *vaniljakastike* = 'vanilla sauce' would initially appear as *vaniljakaeflha*). In the change condition, the second constituent was always a non-word prior to the display change. 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 the second constituent 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.

The target sentences were presented in Courier font (so that each character position was of equal width). The sentences occupied a maximum of 2 lines of text, and the critical word always appeared on the first line but it never was the initial or final word of the line. With a viewing distance of about 60 cm, each character subtended approximately 0.3 degrees of visual angle. There were four blocks of sentences; each member of the matched quadruplet appeared in a separate block. However, each participant only saw two of the four blocks so that he or she saw all the target words only once (i.e., either in the Change or No Change condition). There were 10 items in each condition per participant. The order of the blocks was counterbalanced across participants, and the order of the target sentences was randomised within each block. There were 96 filler sentences among the critical sentences. The text was presented on a 17" ViewSonic (P775) monitor as white against a dark background. The refresh rate of the monitor was set at 150 Hz.

Procedure. The experimental procedure was identical to that of Experiment 1 and 2.

Results and discussion

The experimental design included three variables, transparency (transparent vs. opaque), display change (change vs. no change), and first constituent frequency (low vs. high) that were all within-participant variables and all except frequency were within-item variables. In the display change condition, 21.4% of the trials were excluded from the analyses due to the change taking place after a fixation had already started on the second constituent (these were trials where the eyes crossed the invisible boundary just before the end of the critical saccade).[5] We start by reporting the gaze duration, followed by more detailed analyses of the eye movement pattern. The means of the eye movement measures are given in Table 4.

Gaze duration. There were reliable effects of both first constituent frequency and display change on gaze duration. Gaze durations on compounds with a low-frequency initial constituent were 64 ms longer than gazes on compounds with a high-frequency initial constituent, $F_1(1, 22) = 47.78$, $p < .001$, $F_2(1, 38) = 5.01$, $p < .05$, and gazes were 105 ms longer when there was a display change than when there was no display change, $F_1(1, 22) = 59.49$, $p < .001$, $F_2(1, 38) = 77.23$, $p < .001$. However, averaged over the other two variables, the mean gaze duration on transparent compounds was actually 6 ms greater than the mean gaze duration on opaque compounds (Fs < 1). There was, as in Experiment 2 but not as in Experiment 1, a somewhat bigger frequency effect for the opaque compounds (71 ms vs. 59 ms). However, this interaction, as well as all the other interactions of transparency with the other two variables were far from significant (Fs < 1). Moreover, there was no evidence in the gaze duration measure to suggest that opaque compounds were more strongly affected by the display change than transparent compounds: the effect of the display change manipulation was 104 ms for transparent compounds and 107 ms for opaque compounds. There was a suggestion that the display change effect was slightly greater for compounds with low-frequency first constituents than for those with high-frequency first constituents (114 ms vs. 96 ms), but, as indicated above, this difference was not close to significant. This robust display change effect indicates that readers did process the orthographic information of the second constituent while fixating on the first constituent and subsequently utilized it regardless of the transparency of the compound word and the frequency of its first constituent. The effect size is very similar to that observed by Hyönä *et al.* (in press).

[5] The percentages of excluded trials were 17.0, 20.9, 27.0, and 23.5, for the transparent high-frequency first constituent, opaque high-frequency first constituent, transparent low-frequency first constituent, and opaque low-frequency first constituent condition, respectively.

TABLE 4

Experiment 3: Various reading indexes on the target word as a function of semantic transparency, first constituent frequency and display change

	Transparent compounds				*Opaque compounds*			
	Low-frequency 1st constituent		*High-frequency 1st constituent*		*Low-frequency 1st constituent*		*High-frequency 1st constituent*	
Eye movement measure	*Change*	*No change*	*Change*	*No change*	*Change*	*No change*	*Change*	*No change*
Gaze duration (in ms)	652	535	580	490	649	538	574	472
First fixation duration (in ms)	224	215	203	209	222	208	212	218
Subgaze on the 1st constituent (in ms)	266	272	234	228	271	245	244	235
Subgaze after change	439	339	435	361	461	367	424	359
Probability of refixating 1st constituent	0.19	0.23	0.12	0.07	0.17	0.14	0.1	0.07
Duration of first fixation on the 2nd constituent (in ms)	241	225	248	226	241	217	244	225
Probability of more than 2 fixations	0.59	0.41	0.48	0.35	0.63	0.41	0.52	0.31
Per cent of saccades leaving target word that are regressions	5.3	4.3	5.4	3.9	4.2	5.8	3.5	1.9
Per cent of forward saccades that skip word N + 1	46.1	48	43.8	51.8	49.4	40.5	49.5	45.1
Duration of first fixation on Word N + 1 (in ms)	214	207	224	216	197	196	209	213

More detailed measures on the target word. In order to get a more detailed picture of the time course of the effects, a set of early and later processing measures were analysed. Given that the detailed measures revealed little in the first two experiments, we concentrated on a 'rougher' division corresponding to the display change. That is, one can divide the gaze duration into the time spent on the word before the first saccade over the constituent boundary (i.e., before the display change in the display-change conditions) and the time spent on the word after that saccade. Accordingly, we defined *subgaze$_1$ duration* as the sum of the durations of all fixations that landed on the first constituent prior to the initial saccade to the second constituent (trials on which there was no initial fixation on the first constituent were excluded). The 29 ms main effect of first constituent frequency was nearly significant for this subgaze duration measure, $F_1(1, 22) = 31.46$, $p < .001$, $F_2(1, 35) = 3.33$, $p < .10$. Although this frequency effect was greater for transparent (38 ms) than for opaque (19 ms) compounds, the Frequency $\times$ Transparency interaction was significant only in the participant analysis, $F_1(1, 22) = 4.93$, $p < .05$, $F_2 < 1$. Most strikingly, subgaze$_1$ was only 9 ms greater when there was a display change than when there was not (Fs < 2.1).

A finer analysis of processing before the first fixation on the second constituent revealed little else of interest. The only effect on a more detailed measure that was significant was that the probability of making a second fixation on the first constituent prior to saccading to the second constituent was .18 for the low-frequency and .09 for the high-frequency first constituent targets, $F_1(1, 22) = 33.75$, $p < .001$, $F_2(1, 38) = 6.88$, $p < .01$. This difference accounts for most of the subgaze$_1$ effect above. The only other effect that came close to reaching significance was the Frequency $\times$ Display Change interaction on the first fixation duration, $F_1(1, 22) = 4.16$, $p < .10$, $F_2(1, 35) = 4.19$, $p < .05$, which suggested that the frequency effect was larger in the display change condition than in the control condition.

To summarise, the most striking thing about processing before the constituent boundary is crossed is that there was virtually no early effect of display change in spite of the large effect of this manipulation on gaze duration. In contrast, there were early effects due to the frequency of the first constituent. Although the effect was not quite significant over items for the subgaze$_1$ measure, it was quite reliable in the probability of making a second fixation prior to fixating the second constituent. In addition, consistent with the overall null effects, there were no reliable early effects related to semantic transparency.

As indicated above, we defined a summary measure of later processing which we called the *subgaze$_2$ duration*. It was the sum of the durations of all fixations on the compound word (i.e., on either constituent) after the initial saccade over the constituent boundary (i.e., after the display change in the display change condition). Trials with no fixations on the target word after the display change were excluded from the analysis. Here the only reliable effect was the 83 ms main effect of display change, $F_1(1, 22) = 52.70$, $p < .001$, $F_2(1, 38) = 44.51$, $p < .001$. Thus, the bulk of the overall change effect observed in gaze duration appeared as a relatively late effect (i.e., when the second constituent appeared in its correct form). There was a slight hint in the participant analysis of a Frequency $\times$ Transparency interaction, $F_1(1, 22) = 2.65$, $p > .10$, $F_2 < 1$. Perhaps surprisingly, the first constituent frequency only had a 7 ms effect on the subgaze$_2$ duration (Fs < 1).

The display change effect occurred early during subgaze$_2$, as the duration of the first fixation on the second constituent was 21 ms longer in the display change than in the no change condition, $F_1(1, 22) = 14.81$, $p < .001$, $F_2(1, 38) = 28.37$, $p < .001$. It should be noted that at this point the compound word appeared in its correct form, so the change effect is due to processing done during earlier fixations. In addition, unsurprisingly, the probability of making at least three fixations on the target compound word was affected both by the first constituent frequency, $F_1(1, 22) = 16.08$, $p < .001$, $F_2(1, 38) = 3.54$, $p < .10$, and display change, $F_1(1, 22) = 106.52$, $p < .001$, $F_2(1, 38) = 40.79$, $p < .001$. There were no significant effects on later measures involving transparency.

To summarise the results of later processing measures, two findings came out very clearly. First, the display change effect that was quite absent during the early processing is large and significant in later processing. Second, the first constituent frequency effect that was present in the early measures also had some effects in the later measures, although the overall effect of first constituent frequency on later measures, as measured by the subgaze$_2$ duration, was quite small and not close to significant. In addition, there was little effect of semantic transparency even on later measures.

Measures after leaving the target word ('spillover'). We also examined fixations made immediately after leaving the target word, especially to see if there were any later effects due to transparency. As in Experiment 2, the sentence frames containing the pairs of matched transparent and opaque words were identical up to the word following the target word. Thus, we analysed the duration of first fixation made on the word N+1 as a function of the manipulated variables. One participant was removed from the participant analysis and eight items were removed from the item analysis due to missing data. There was a non-significant trend for a transparency effect, $F_1(1, 21) = 2.52$, $p > .1$, $F_2(1, 30) = 4.42$, $p < .05$, but the trend was 11 ms in the wrong direction (i.e., the transparent condition was associated with a longer fixation duration than the opaque condition). There was a tendency for a frequency effect in the participant analysis, $F_1(1, 21) = 3.63$, $p < .10$, $F_2 < 1$, but again in the wrong direction (i.e., the high-frequency condition was associated on average with a 12 ms longer fixation than the low-frequency condition). It is very likely that the trend reflects the fact that Word N + 1 was not identical between the two frequency conditions.

The probability of leaving the target compound with a regression showed no reliable effects, all Fs < 1.5 (see Table 4). Another measure indexing late processing, the probability of skipping word N + 1 yielded an almost significant Transparency × Display Change interaction, $F_1(1, 22) = 3.95$, $p < .10$, $F_2(1, 38) = 3.88$, $p < .10$. The nature of this marginal interaction is that with opaque compounds, word N + 1 was skipped more often in the change than in the no change condition, whereas an opposite trend was evident for transparent compounds.

GENERAL DISCUSSION

Our data indicate that there was no effect of transparency for the global 'first pass' processing time on a word (i.e., gaze duration). In Experiment 1, there was virtually no main effect of transparency (1 ms) nor any interaction with frequency. In Experiment 2, although there was a suggestion of a transparency effect (a main effect of 11 ms), the pattern of the effect is counter to what would be predicted by most views of how transparency would affect the processing of compound words. That is, there was virtually no transparency effect for the words with high-frequency first constituents, and the entire transparency effect occurred for the words with low frequency first constituents. If composition of the meanings of the constituents was an important part of recognising a compound word, one would have expected that this process would be particularly important when the first constituent was frequent, as the compositional process would have a head start over a direct look-up process. Moreover, in Experiment 3, the transparency main effect on gaze duration was actually 6 ms in the 'wrong direction' (i.e., opaque words having shorter gaze durations than transparent ones).

In addition, the more detailed analyses of 'first pass' processing also provided no convincing evidence for transparency effects. There were a few suggestions of transparency effects on other measures, but they were not consistent across experiments nor were they in places where one would expect to find them. Most notably, the effect of transparency on the first fixation duration in Experiment 1 was not replicated in either Experiment 2 or Experiment 3. As mentioned earlier, the comparison in Experiment 1 was across different sentence frames (and thus may have been due to a confounding of the difficulty of the sentence frames with transparency) whereas the designs in Experiments 2 and 3 did not have this confounding.

Moreover, the duration of the first fixation would seem like a strange place to find a transparency effect, as one would expect transparency to only affect later processing (e.g., second and third fixation durations) — presumably only after the meaning of the second constituent had been identified. This argument is bolstered by the finding that the display change manipulation in Experiment 3 (where the second constituent was either visible at the outset or only when it was fixated) had no effect on first fixation duration or other measures of early processing. In Experiment 2, where main effects of transparency would be picked up more reliably than in Experiment 1, there was a suggestion of a transparency effect on the third fixation duration—but a suggestion of an opposite effect on the duration of the second fixation. There were also four transparency by frequency effects that were reliable or marginally reliable over participants but not close to reliable over items, all in the direction of a greater frequency effect for transparent compounds. However, their locus was not consistent: in Experiment 2, these interactions were found in second and third fixation duration, whereas in Experiment 3, they were on first fixation duration and subgaze_1 duration (the latter two measures are not independent). Thus, besides being not reliable over items, these effects were not consistent across experiments and were opposite to the direction of the (non-significant) interaction observed in the gaze duration. Moreover, as indicated above, early measures are implausible places to find a transparency effect, as it is unlikely that the meaning of the second constituent is processed quickly. Thus, we think that the best conclusion one can draw from the three experiments is that there is no transparency effect on 'first-pass time' for processing Finnish compound words, or if there is an effect, it is so small to be of no real importance in understanding how compound words are processed in context.[6]

We should also comment briefly on the results of our display change manipulation, although that was not the primary focus of the present study. Essentially, these results replicated those of Hyönä et al. (in press) in finding: (a) a strong effect of the first morpheme frequency before the second constituent is fixated, but virtually no effect of the display change; and (b) a strong effect of the display change manipulation after the second constituent is fixated, but only weak effects of the first constituent frequency on these later measures.[7] This pattern of results suggests that

[6] It is possible that the processing of opaque compounds receives benefit from a sentence context (even a neutral one).

[7] This points to another reason that it is unlikely that transparency would affect early measures. That is, if early measures do not even pick up whether a meaningful second constituent is present or not, they are quite unlikely to be sensitive to whether the composition of the meanings computes semantically.

these long Finnish compounds are processed quite serially, with the initial fixations reflecting processing of the first constituent and the later fixations reflecting processing of the second constituent. They also indicate that constituents yet to be fixated (i.e., material in the parafovea) have little influence on the duration of the initial fixation.

We think it is fair to say that our data indicate that the transparency has little or no effect on the processing of long compound words in reading Finnish and that the compositional processes exposed by manipulating the frequency of the constituents involve 'glueing together' something like the constituent lexemes rather than composing the meanings of those lexemes. The one caveat we need to make is that it is possible that there are 'later' transparency effects that our paradigm failed to pick up. That is, our sentences were only matched up to the word after the target word. Thus it is possible that there might be more later regressions back to the target word for opaque words, indicating some sort of 'double take'. Analyzing such later measures with our materials would be fruitless, as the putative transparency effects would be hopelessly confounded with differences in the sentence frames. However, our design did permit analysis of some earlier 'spillover' effects, and again, there was little or no effect of transparency on these measures.

Another possible caution we should mention in interpreting our data is that the pattern of results may be restricted to Finnish compound words. However, we think it is quite unlikely that there is anything special about Finnish that would produce a reduced effect of transparency. In fact, most plausibly the opposite is true, as Finnish compounding is very productive. As a result, Finnish readers frequently need to put together the meanings of constituents of novel compound words in order to understand them. Thus, the fact that transparency has virtually no effect on the understanding of already existing compounds is indeed quite remarkable. It suggests that the composition process may exist in two stages: first a composition of entities that does not involve composing the meanings followed by a composition of meanings, if that is needed for novel compounds.

Basically, our results are consistent with most of the work using the priming techniques we reviewed earlier. That is, a large majority of the lexical decision experiments that used variations on the constituent priming paradigm (i.e., priming the compound with a constituent or vice versa) failed to obtain a transparency effect either. However, in one of these studies and in two of the semantic priming experiments we reviewed, there was an indication that there might have been a smaller priming effect for completely opaque words such as *cocktail*, where the relation of the parts to the whole would be known only to specialists of the language. Our opaque words were a mixture of such completely opaque words and those that were partially opaque (i.e., where the meaning of one of the constituents was related to the meaning of the compound). Given our overall null results, we were sceptical that the relative opaqueness would make any difference. Nonetheless, we decided to perform a post-hoc item analysis where we predicted the difference in gaze duration between each matched transparent-opaque target pair of Experiment 2 by the relative difference in the transparency rating of the target pair. In this analysis, the correlation between the difference in transparency rating with gaze duration was small ($r = -.137$, $F < 1$) and was in the 'wrong' direction; a bigger difference in rated transparency between the transparent and opaque words predicted a somewhat smaller difference in gaze duration.

What kind of processing model could explain the observed pattern of data? There seem to be two types of model that would be plausible. The first is the one that we have implicitly been assuming. That is, the lexical entries of the components are accessed, which then are glued together to 'look up' the lexical entry of the whole word. For example, in English, one would first access the lexical entry for 'straw' and then the entry for 'berry', which in turn would look up a lexical entry for 'straw–berry' from which the meaning of the compound would be looked up. Such a componential process is likely to be going on in parallel with a direct look-up process, where the entire string of letters is accessing 'strawberry' directly. These two processes could either be independent processes or could be working interactively. Either version of such a dual-route model would be consistent, at least qualitatively, with both constituent frequency effects and whole-word frequency effects.

The second type of model is something like Taft and Forster's (1976) two-stage model, where the first constituent accesses a first stage 'file drawer' that contains all the entries that have the initial constituent as a first constituent. The second stage of the access would then be 'finding' the compound word in the 'file drawer'. However, the second stage needs not be serial, as Taft and Forster originally posited, but, as in Lukatela, Gligorijevic, Kostic, and Turvey's (1980) 'satellite-entries' model, could be a parallel process which is determined by both the number of competitors in the file drawer and/or their frequency relative to the frequency of the word actually seen. Such a model is also consistent with both first constituent frequency effects (the ease of finding the 'file drawer') and whole word frequency effects (whole word frequency will be confounded with the position of the item in the file drawer assuming first constituent frequency has been equated).

What do the two types of models predict about transparency effects? For the file drawer model, the meanings of the two constituents seem besides the point; the access of the meaning is through the whole-word entry. In contrast, the dual route model could make any prediction, depending on whether the components that are activated and combined are ‘lexical’ or ‘semantic’’. Thus, the present results are consistent either with a two-stage model, such as Taft and Forster (1976), or with a dual-route model in which there is no significant activation of meaning by the access of the constituents in the combination process. However, a critic may argue that possible processing differences between transparent and opaque compounds may be obscured by differences in the number of competitors in the file-drawer (the size of the morphological family) and/or the relative ranking of the transparent and opaque compounds in the file-drawer. To examine this possibility, we did some post-hoc analyses to determine whether these variables significantly affected processing in the present experiments.

We computed a measure of the morphological family size for each target word by counting the number of compound words that existed in our computerised corpus (Laine & Virtanen, 1999) given the first constituent (possible allomorphic variation was taken into consideration; the so-called positional family size, see De Jong, Feldman, Schreuder, Pastizzo, and Baayen, 2002). We also computed the relative ranking of each target word in the family. As expected, the family size was clearly bigger for the compounds in the high-frequency than in the low-frequency first constituent condition. It was also somewhat bigger for transparent than opaque compounds. The mean sizes were 314, 218, 48, and 27 for the high-frequency transparent, high-frequency opaque, low-frequency transparent, and low-frequency opaque conditions, respectively. As one might expect, the relative ranking in the family was higher in the low-frequency than in the high-frequency condition. It was somewhat higher for transparent than opaque compounds. The mean rankings were 27.3, 15.3, 2.1, and 1.4 for the high-frequency transparent, high-frequency opaque, low-frequency transparent, and low-frequency opaque conditions, respectively. In the low-frequency conditions, the ranking was 1 or 2 for all but five words.

The prior data suggest that a big family facilitates compound word processing (De Jong et al., 2002). This would imply that our present confounding of transparency with family size should have produced a transparency effect. In contrast, the prior data suggest that a high ranking within the family facilitates processing; this confounding with transparency should work against a transparency effect. Thus, it is not obvious that there is a confounding problem. However, we did post-hoc correlational analyses on the item means for each of the experimental conditions of Experiment 2 and 3 to examine whether family size and relative ranking within the family had any effect on the pattern of data. We first computed the bivariate correlation of family size and ranking, and the correlation of each with gaze duration and the two sub-gaze measures and found that (a) family size and ranking correlated quite highly, and (b) family size produced stronger correlations with the eye fixation measures than ranking. As a result, in the subsequent analyses, we used family size to predict the fixation time measures. As word frequency and first constituent length might be confounded with family size in these analyses, these variables were also entered as predictors.

The multiple regression analyses demonstrated that for all other conditions except for the opaque low-frequency first constituent condition the relationship between family size and gaze duration is depicted by a small negative slope that was non-significant. In other words, the bigger the family the shorter the gaze duration is. However, for the opaque compounds having a low-frequency first constituent the slope was clearly positive. The slope was statistically significant ($p < .05$) for the gaze duration of both the no-change and the change condition of Experiment 3; for the gaze duration of Experiment 2, the slope was positive, but did not reach significance. For these opaque compounds, there appears to be an inhibition effect due to family size: the bigger the morphological family, the longer the gaze duration is. To examine the time course of this inhibition effect, we then computed the regression analyses on the two sub-gaze measures of Experiment 3 (separately for the no-change and the change conditions). In Subgaze_1 (i.e., the summed duration of first-pass fixations on the first constituent) there was no indication of an early inhibition effect, whereas Subgaze_2 yielded a significant positive slope for family size in both the change and no change conditions, indicating that the effect appears relatively late.

These regression analyses thus suggest that, although there is no evidence that opaqueness exerts any overall effect of slowing identification of compound words, opaqueness may play some role in compound word processing if the frequency of the initial constituent is relatively low. We are not sure why opaqueness played a role only for the compounds with low first constituent frequency, however. Here is one speculation. As the frequency of the compound word was held constant across conditions, the low-frequency first constituent conditions had a much smaller family size (see above) and also the compound word was more often close to the most frequent member of the family. In a recent paper (Hyönä et al., in press) we found suggestive evidence that there was a late facilitation effect suggestive of a predictability effect on processing the second constituent when the family size was low. What may be critical is whether the word that is seen is the most frequent member of the family or not. That is, if the observed word is the most frequent, it may be ‘predicted’ and thus there would be little interference, whereas if it is not, some other compound word (which is presumably usually transparent) will be predicted, and its activation will cause interference with obtaining the correct meaning for an opaque compound when the word is ultimately identified. If this analysis is correct, it suggests that there may be transparency effects, but for a very limited set of compound words. Of course, one should treat these analyses with some caution, as they were post-hoc and thus there could be some other variable confounded with family size that was the real underlying variable.

One question that our experiments leaves open is whether the entity that starts either of these hypothetical compositional processes needs to be a constituent or a morpheme. For example, perhaps any reasonable syllable or common orthographic pattern will do. Although there may be different answers in different languages, there is evidence in Finnish that morphemically defined constituents have a privileged status. Laine, Vainio, and Hyönä (1999) observed no difference in lexical decision times between pseudo-inflected words and their monomorphemic controls, whereas inflected words consistently produced longer decision times than their monomorphemic controls. These data may be taken as evidence against the view that all letter clusters within a word that can potentially form a morpheme would be automatically activated. Further corroborative evidence against this view comes from an eye-tracking study of Bertram, Pollatsek, and Hyönä (in press) on compound word processing. In Experiment 1, there was a set of compounds that contained a pseudomorpheme within a two-constituent compound. The identification of these compounds took no longer than for compounds that did not contain a pseudomorpheme. Moreover, the finding reported by Bertram and Hyönä (2003) that there was no constituent frequency effect for short compounds (7–9 characters) indicates that the constituent frequency effect is unlikely to be a general orthographic effect. If it were, the constituent frequency effect should show up for short and long compounds alike.

Some of the priming studies discussed in the Introduction also have addressed the question about the nature of the constituent effect observed for opaque and transparent compounds. Zwitserlood (1994) was able to rule out an orthographic explanation, as she failed to find a priming effect for a letter string that was not a compound word constituent (in her orthographic prime condition of Experiment 1, the prime was a compound and the target was a word or a pseudoword; *kerstfeest* → *kers*). The masked priming study of Longtin, Segui, and Hallé (2003) observed a priming effect of similar magnitude for semantically transparent, opaque and pseudo-derived words (not compounds), whereas no priming was obtained for orthographic controls. This study points to an early parsing mechanism that is blind to semantic transparency and blind to whether or not the two morphemes actually combine to form an existing complex word. This seems consistent with Monsell's (1985) constituent priming results, as he found that there was no significant difference in priming for pseudocompounds (*boy* → *boycott*) and priming for real compounds (although the priming effect was somewhat smaller for pseudo-compounds than for real compounds).

The bulk of the existing evidence thus points to a morphological interpretation for the constituent frequency effect observed in the present study. Hence, we suggest that, in reading text, lexical entries for the constituents of both transparent and opaque long compound words are used in a compositional mechanism that is part of the word encoding process. However, the lack of a transparency effect indicates that the composition of meaning is not an important part of this process, at least in the initial identification of the compound word.

REFERENCES

Andrews, S. Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in the mousetrap. *European Journal of Cognitive Psychology*, *16*, 285–311.

Bertram, R., & Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory and Language*, *48*, 614–634.

Bertram, R., Pollatsek, A., & Hyönä, J. (in press). Morphological parsing and the use of segmentation cues in reading Finnish compounds. *Journal of Memory and Language*.

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 and Language*, *81*, 555–567.

Hyönä, J., Bertram, R., & Pollatsek, A. (in press). Are long compound words identified serially via their constituents? Evidence from an eye-movement contingent display change study. *Memory and Cognition*.

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

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

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.

Laine, M., Vainio, S., & Hyönä, J. (1999). Lexical access routes to nouns in a morphologically rich language. *Journal of Memory and Language*, *40*, 109–135.

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

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

Libben, G., Gibson, M., Yoon, Y., & Sandra, D. (1997). Semantic transparency and compound fracture. *CLASNET Working Papers*, *9*, 1–13.

Libben, G., Gibson, M., Yoon, Y. B., & Sandra, D. (2003). Compound fracture: The role of semantic transparency and morphological headedness. *Brain and Language*, *84*, 50–64.

Longtin, C.-M., Segui, J., & Hallé, P. A. (2003). Morphological priming without morphological relationship. *Language and Cognitive Processes*, *18*, 313–334.

Lukatela, G., Gligorijevic, B., Kostic, A., & Turvey, M. T. (1980). Representation of inflected nouns in the internal lexicon. *Memory and Cognition*, *8*, 415–423.

Monsell, S. (1984). Repetition and the lexicon. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol. 2). Hove, UK: Lawrence Erlbaum Associates Ltd.

Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 820–833.

Pollatsek, A., Reichle, E., & Rayner, K. (2003). Modeling eye movements in reading. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 361–390). Amsterdam: Elsevier.

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

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., Burani, C., & Baayen, R. H. (2003). Parsing and semantic opacity. In E. M. H. Assink & D. Sandra (Eds.), *Reading complex words: Cross-linguistic perspectives* (pp. 159–189). Dordrecht: Kluwer Academic.

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

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