

What we look at when we read'

J. MEHLER, MASSACHUSETTS INSTITUTE OF TECHNOLOGY

T. G. BEVER, MASSACHUSETTS INSTITUTE OF TECHNOLOGY AND HARVARD UNIVERSITY. P. CAREY, HARVARD UNIVERSITY

The entire surface phrase structure hierarchy of sentences influences visual scanning patterns in reading familiar material. A general eye-fixation rule predicts the pattern of adults' eye-fixations in a study of reading predictable sentences: "fixate on the first half of phrase structure constituents." This rule applies cumulatively at all levels of the surface phrase structure.

The psychological nature of what has just been perceived influences the ongoing pattern of search for new information. The study of eye-fixations in response to visual stimuli such as written sentences offers precise data on the interaction between stimulus configurations and search patterns: Visual stimuli are easy to vary, and the movements of the eyes may be taken as an external reflection of the sensory search pattern.²

Transformational grammar provides for each sentence a unique structural description. Experimental studies have shown that many of these linguistic structures interact with short- and long-term memory, comprehension of language, and the learning of language (Chomsky, 1965; Miller, 1962; Mehler & Savin, in press). In the research reported in this paper we take this linguistic theory as the prior basis for the structural analysis of the sentences and we investigate the question: Which aspects of sentence structure affect the visual scanning pattern in reading?

Transformational grammar specifies for each sentence several phrase structure hierarchies, which represent the groupings of adjacent words and phrases into "immediate constituents." For instance in sentence (a) below, "dog" is less closely bound to "candies" than it is to "his" while in sentence (b) "dog" is more closely bound to "candies" than to "him."

(a) They fed his dog candies
(They) ((fed) ((his) (dog) (candies)))

(b) They fed him dog candies
(They) ((fed) (him) ((dog) (candies)))

(The conventional linguistic notation is indicated by the bracketed sequences. Each pair of a left and right bracket corresponds to an "immediate constituent.") Every single word is itself an immediate constituent. Immediate constituents occur in a hierarchy, so that one word (e.g., "dog") can be simultaneously contained within several constituents. In sentence (a), in addition to the separate words, the constituents are:

They fed his dog candies

fed his dog candies

his dog

"Dog" is contained in three constituents. This is

also the case for sentence (b), but the last constituent is different from that of (a).

They fed him dog candies

fed him dog candies

dog candies

The following pair of sentences also differ structurally in the arrangement of their immediate constituents.

(c) They are smiling authors

(They) (are) ((smiling) (authors))

(d) They are feeding authors

(They) ((are) (feeding)) (authors)

These relations are represented in the phrase structure assigned to each sentence by the grammar. Transformational grammars express two easily defined levels of phrase structure relations. The most obvious level of phrase structure is at the "surface structure" as in the above examples; different surface structures are easily noticed since they effect differences in the prosodic features of pronunciation (e.g., stress, pitch, rhythm). Consider, for example, the different stress and rhythm patterns associated with the different surface phrase structures in (a) and (b), and (c) and (d).

The level of the "deep phrase structure" specifies an additional unique set of phrase structure relations for each sentence. This structural level represents the "logical relations" which the words and phrases bear to each other. The deep phrase structure specifies the logical "actor" and "action" (and "object" if the action is transitive) independently of the actual order of the words. For example, in (c) "authors" is the deep phrase structure actor of the preceding verb, while in (d) it is the object. Often there are differences in deep phrase structure which have no direct reflection in surface structure or in pronunciation: The two sentences below have identical apparent phrase structure but a marked difference in the logical relations, since "hunters" is the subject of "stare" in (e), but the object in (f).

(e) The staring of the hunters was awful

(f) The staring at the hunters was awful

This study treats the following questions: 1) Does phrase structure affect eye-fixation patterns in reading? 2) If so, do surface phrase structures and deep phrase structure have an equal effect? 3) What is the descriptive calculus which describes precisely the interactions between grammatical structures and eye-fixation patterns?

METHOD

Materials

The experiment was not concerned with the effects on eye-fixation patterns of changes in vocabulary or

spelling. To avoid these lexical differences which appear in the above three pairs of sentences, three corresponding types of ambiguous phrases were constructed. Each ambiguous phrase has two interpretations which differ structurally in the same way as the sentence pairs above. The interpretations of each ambiguous phrase differ mainly at the surface phrase structure level (SS) (like the structural difference between (a), (b)), at both levels (SS, DS) (like (c), (d)), or just at the deep phrase structure level (DS) (like (e), (f)). There are two examples of each type of ambiguous sentence, matched for number of words and length so that each subject could be exposed to one example of each type of grammatical structure without repeating any lexical sequences.³

- (SS) 1. they gave her dog candies. . . .
 2. they told her cat stories. . . .
- (DS,SS) 1. they are surprising authors. . . .
 2. they are interesting farmers. . . .
- (DS) 1. the shooting of the hunters. . . .
 2. the punching of the sisters. . . .

Each (SS) sentence has two interpretations which differ mainly in the surface phrase structure grouping of adjacent words. For example, SS-1 has either a surface phrase structure like sentence (a) or like (b) (see above), depending on its interpretation.

- SS-1 (a) they (gave (her dog) candies). . . .(that is candies were given to her dog)
 (b) they (gave her (dog candies)). . . .(that is dog candies were given to her)

DS-SS sentences have either a surface phrase like sentence (c) or like (d).

- SS-DS-1 (c) they (are (surprising authors)). . . .(that is "those authors are surprising to people")
 (d) they ((are surprising) authors). . . .(that is "somebody is surprising the authors")

In addition, the underlying phrase structure subject/object relations are reversed: in interpretation SS-DS-1-(c), "authors" is the underlying subject of "surprise" (since it is the authors who are doing the surprising). In interpretation SS-DS-1-(d), "authors" is the underlying object of "surprise" (since somebody else is surprising the authors). Finally, the two interpretations of DS-1 have the same surface phrase structure, just as sentences (e) and (f) below.

- DS-1 (e) (the shooting (of (the hunters))). . . .(that is "the hunters were shooting")
 (f) (the shooting (of (the hunters))). . . .(that is "somebody was shooting the hunters")

The underlying structures of DS-1-(e) and DS-1-(f) differ in the subject/object relations since in DS-1-(e) "hunters" is the underlying subject of "shoot," while in DS-1-(f) it is the underlying object.

The semantic interpretation of these sentences can determine the particular structural interpretation which they receive. Therefore, each of these phrases was embedded in the fourth sentence of a five-sentence story which made one interpretation highly probable. In SS-1, for instance, the sentence was either preceded by three sentences about a girl who ate dog candies, or about a girl's dog that liked chocolate candies. Twelve series of five sentences were constructed, one series for each ambiguous phrase with one of its interpretations. Since there were two examples of each type of ambiguous phrase, each subject read each kind of sentence structure once without repeating any lexical sequence (six structures in all). None of the sentences were punctuated except for a single blank space between words. All sentences had 37-40 letters and 8-10 words. Each sentence was on a separate black-on-white slide.

Procedure

Forty paid Harvard and Radcliffe students participated in individual sessions of approximately 45 min. each. Subjects (Ss) sat before an optical apparatus developed by N. Mackworth for recording eye-fixations.² This device photographs a beam of light reflected from the S's left eyeball. As the eye moves, the angle of reflection changes and the reflected light beam sweeps across a photographic field. Whenever the eye movement stops briefly for a fixation (of at least 0.1 sec.), the light beam rests in the same part of the photographic field and overexposes a spot on the film. The visual stimulus that the S sees is superimposed on the film at the same time. So in our experiment a response protocol consisted of a photograph of a sentence with bright dots indicating where in the sentence the S fixated for a short time.

In each experimental session, E projected an alignment sentence on the stimulus screen. E instructed S that seven series of five sentences would be presented, one sentence at a time. (There was one practice series, and six experimental series for each S.) He was to read each sentence once, from left to right, then fixate on the right hand edge of the screen until the next sentence appeared. In each series a new sentence appeared every 5 sec. S was told that each series formed a story and that on completion he would be requested to retell the story aloud. Each sentence was about 15 in. from S's right eye, and subtended an angle of 16 degrees. The apparatus was calibrated for each S before and after each series to make sure that the stimulus sentence and reflected light-beam were correctly superimposed.

Experimental data consisted of pictures of S's fixation points superimposed on the ambiguous sentence during reading. Although Ss were instructed to move as little as possible, small shifts in head position often occurred during a series of five sentences. Unfortunately, these movements displaced the alignment of the sentence and the reflected light beam.

Table 1. Number of Eye Fixations Summarized by Sentence Structure, Normalized to 10 Ss for Each Individual Sentence.

		gave told		(her her)		(dog cat)		and stor	ies ies	Actual No. Ss Contributing Data	Actual No. Eye Fixations	
SS	(a)	14.3	.35	13.7	2.6	8.7	2.1	10.3	9.8	24	73	
SS	(b)	13.5	3.6	7.9	3.6	14.3	2.8	10.4	10.7	21	70	
SS-DS	(c)	10.7	1.6	12.6	7.3	3.2	4.2	12.6	8.2	19	54	
SS-DS	(d)	9.8	2.8	8.3	13.9	5.3	2.7	15.6	9.1	18	61	
DS	(e)	14.7	5.8	2.2	6.9	4.1	9.5	2.2	14.3	6.4	19	63
DS	(f)	12.7	7.2	2.5	4.6	2.0	8.1	3.9	12.7	5.8	22	64

Records in which accidental displacements of less than four letters to right or left had occurred during the series were retained, and all fixation spots in those records were adjusted by the amount of displacement. Records were retained with the further requirement that S recalled only the appropriate interpretation of the ambiguous sentence. (117 records out of 240 were discarded because they did not meet one or the other of these two criteria.)

RESULTS

The 40 Ss provided 123 records which were satisfactory. These were scored to the nearest half-letter (after adjustment for any overall displacement) since the recording system did not justify more precise analysis. Due to the poor quality of particular records, each of the six sentences with its two different structural interpretations had different numbers of contributing Ss. The range was from seven to 14 Ss for each distinct lexical sequence with a particular interpretation. The data for each sentence interpretation were normalized to ten Ss each, so that the sentences could be directly compared. Table 1 summarizes the normalized eye-fixation patterns for each of the sentence-structure combinations. For purposes of analysis, each of the sentences was divided into its morphemes (where possible) and spaces. The total number of fixations on the two morphemes which have the same syntactic structure are summed in the numbers in each row in Table 1. For example, in SS, the total number of eye-fixations occurring somewhere in the words "dog" and "cat" with interpretation (a) was 8.7: In interpretation (b) it was 14.3 (each of the six summaries is the result of two

separate lexical sequences with the same syntactic structure).

Figure 1 summarizes the eye-fixation pattern differences between the two interpretations for each type of structurally ambiguous sequence. The three graphs of Fig. 1 are derived from the three pairs of summaries in Table 1. For example, graph SS is the result of the subtraction of each value in SS (b) from the value of SS (a) for the same sequence of letters and spaces. For example, in Graph SS, the value for *dog & cat* is -5.6 since that is the result of subtracting 14.3 from 8.7. In this way Fig. 1 expresses the effect on eye-fixation patterns of each of the three kinds of differences in syntactic structure.

Examination of Fig. 1 indicates that the surface structure differences (in SS and SS+DS) were reflected by noticeable differences in the eye-fixation patterns. The structure which differed only at the deep phrase structure (DS) did not show such large differences. To quantify this comparison, we calculated the average difference per letter correlated with each type of structural change. For each sentence we divided the total of the absolute values of the differences in Fig. 1 by the number of letters in each sentence. For example, the total value of the absolute differences in eye-fixations between the two structural interpretations of SS sentences is 18.25. The number of letters and spaces is 20, so the difference per letter brought out by the two structural interpretations in SS is .91. The average differences per letter in the two interpretations of phrases of type SS+DS is .91, and for type DS, .58.

For purposes of comparison we calculated the amount of difference between sentences which main-

tained structure, but changed vocabulary (e.g., "they gave her dog candies," and "they told her cat stories," both with the structural interpretation in which "dog candies" and "cat stories" are constituents). The result of this was an average difference of .74 per letter for all lexical comparisons: Thus, changing the surface phrase structure results in a larger change than changing the vocabulary and holding the structure constant, while changing the deep phrase structure has a smaller effect than changing the vocabulary. This indicates that surface phrase structure is the linguistically defined level which interacts most strongly with the visual scanning pattern.⁴

DISCUSSION

The exact nature of the interaction between the surface phrase structure and visual scanning is not obvious from the few examples which we have, but certain facts are extremely indicative of the general pattern. Consider first the surface structure difference between single words like "farmers" and "hunters" on the one hand and "authors" and "sisters" on the other hand. The first two words have a complex structure since they are composed of the words "farm" and "hunt" with a plural agentive suffix added. Thus, they have the structure indicated in (j) while the other two words have no internal structural analysis:

- (j) ((farm) ers), ((hunt) ers)
- (k) (authors), (sisters)

These two pairs of words occurred in structurally identical contexts during the experiment so that they

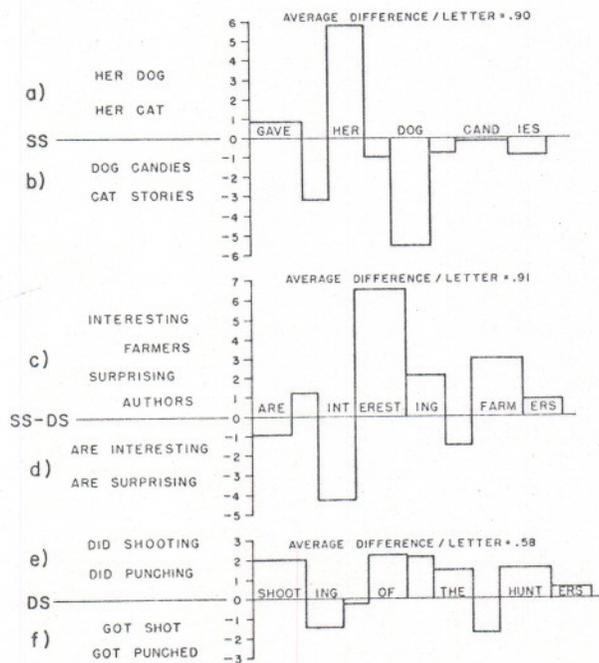


Fig. 1. Differences in eye-movement patterns correlated with differences in symbolic structure. See text.

Table 2.

farm	ers
1	(fixate on beginning of "farm")
1	(fixate on beginning of "farmers")
—	
2	total
authors	
1	(fixate on beginning of "authors")
1	total

may be compared for their eye-movement patterns. The ratio of eye-fixation placements on the first and second parts of the words were found to be:

"farmers" "authors"
 "hunters" "sisters"
 3/1 1.3/1

These results indicate that it is the constituent itself that attracts eye-fixations. Suppose we assume the following principle of eye-fixation for reading: *fixate on the beginning of a surface phrase structure constituent.* This principle predicts the above patterns. For "farmers" there are two constituents, one contained within the other: *farm* and *farmers*. The first half of each of them is contained within the letters "far." Table 2 presents the method of calculating the predictions. Each constituent is assigned a fixation on its first half. If a sequence of letters (e.g., "far") begin *n* constituents simultaneously, then it is assigned *n* fixations.^{5,6} Thus, we would predict relatively more fixations on the beginning of "farmers" and "hunters" than the beginnings of "authors" and "sisters."

The above system is sufficient for the description of different eye-fixation patterns on individual words that differ in surface phrase structure. In this experiment we also had two pairs of sentences in which each sentence has two possible surface phrase structures. We have already shown that the surface phrase structure differences are reflected by different eye-fixation patterns. The same sort of descriptive systems as used for the individual words predict the direction of the fixation pattern differences brought out in response to the different surface phrase structures.

If the fixation rule is made more precise, eye fixation patterns can be predicted for each surface phrase structure: The rule is *fixate on the first half of each constituent.* As in the case of single words, this rule applies recursively at all levels of the surface phrase structure. For example, in the structure DS-1-(a) "they told (her cat) stories" the rule operates on each word, as well as on the constituents, *her cat*, and *told her cat stories*. The fixation rule distributes one fixation evenly over the first half of each constituent. The ratios of predicted fixations for SS-(a) structures and how they are calculated are shown in Table 3. In SS-(a) sentences the constituent "(her cat)" is seven letters long, so the fixation is distributed evenly over the first 3.5 letters: This assigns 3/3.5 or .86 of a fixation to the word "her" since it

Table 3. Predicted Ratios of Eye Fixations for SS-a Structures

they	told		(her	cat)	1	stor	ies	
	1		1	1	1	1		(fixate the first half of each word)
			.86	.14				(fixate the first half of the IC "her cat")
	.4	.1	.3	.1	.1			(fixate the first half of the IC "told her cat stories")
	1.4	.1	2.2	.2	1.1	0	1.0	predicted values
	1.4		1.8		1.1		1.0	observed data
								(ratios of eye fixations)

is three letters long, and the remainder, .14, is assigned to the space. Similarly, the constituent "told (her cat) stories" is 20 letters long, so one fixation is distributed over the first 10 letters; "told" receives 4/10 fixations since it is four letters long, "her" receives 3/10 fixation, and so on.⁷ The predictions and observed ratios for each structural interpretation of sentences SS and SS+DS are summarized in Table 4. In general, the relative number on each syllable is predicted by the fixation principle if the predictions are tested for each structure.

The fixation principle also predicts the relative differences for each pair of structures shown in Fig. 1. There are seven cases in SS and SS+DS sentences for which the fixation principle predicts that one of the structures will receive more eye-fixations than the other corresponding structure. There are also four cases in which the fixation principle predicts no difference for the structure. For example, the word "her" in SS-(a) has a prediction of .9 fixations more than in structure SS-(b) (see Table 1: in SS-(a), 2.2 fixations are predicted, in SS-(b), 1.3, giving the difference of +.9). All the predicted differences associated with the differences in phrase structure in SS and SS+DS sentences are shown in Table 5. We tested the correlation⁸ of the size and direction of all 11 predicted differences against the actual differences and found a correlation of $r = .84$. As a further check on the uniqueness of the predictions made by the fixation rule, we calculated the predictions if the fixation rule were "fixate on the last half of the constituent." These predictions were

tested with the results in the same way as above and have an $r = -.54$.

These results show that the actual data are best predicted by the principle as stated above. The exact form of the fixation rule is not clear. There is no doubt that it is influenced by the familiarity of the individual phrases in the sentences, and the experimental task. In everyday reading, for example, we make far less than one fixation every .26 letters, nor do we necessarily read from left-to-right. Furthermore, the fixation function may be a right-skewed curve spread over the entire phrase structure constituent with an emphasis on the first part, instead of a discrete rule which totally ignores the second half of the constituents. Although further investigation of these variables is required, one fact is strongly supported by our results: The fixation strategy is dependent on all surface phrase structure levels of sentences.

In this respect the fixation rule is similar to many "cyclic" rules which have been developed in linguistic theory and for other psycholinguistic phenomena;⁹ these also apply to successively larger phrase structure constituents. Although the data do indicate that all surface phrase structure constituents are involved in this reading situation, it is impossible to tell if the fixation rule also applies cyclically—from the smallest constituent to the largest—as in linguistic analyses, or if the order of application is some other.

The surface phrase structure interacts with many auditory perceptual phenomena in addition to the perception of stress. For instance, clicks presented

Table 4. Predicted and Actual Eye Fixation Ratios for SS and SS-DS Structures

SS	predicted	told	(her	cat)	stor	ies...	told	her	(cat	stor	ies)...		
	observed (ratios)	1.4	2.2	1.1	1.0	0	1.4	1.3	1.6	1.3	0		
SS-DS	predic	(are	int/	erest/	ing)	farm	ers...	are	(int/	erest/	ing	farm	ers)...
	observed (ratios)	1.4	1.8+	.6	.0	1.5	0	1	1.8-	1.1	.1	1.5	0

Table 5. Predicted and Observed Eye-fixation Pattern Differences for lexical items Correlated with Differences in Surface Phrase Structure (Predicted values are multiplied by ten to bring them into the same range on the observed values) (see Fig. 1)

	(SS a) - (SS b)					(SS-DS-c) - (SS-DS-d)					
	told	(her cat)	stor	ies		are	(int	erest	ing	farm	ers)
predicted	0	9	-5	-3	0	-4	-.1	5	.1	0	0
observed (corrected to 10 subjects/stimulus sentence)	.8	5.8	-5.6	-.1	-1.0	-.9	-4.3	6.6	2.1	3.0	.9

during sentences are perceived as occurring between constituents even when they actually occur during constituents (Fodor & Bever, 1965; Garrett, in press). Thus the scanning strategy for reading which we have suggested is a visual analogue of auditory strategies used in the perceptual analysis of spoken language. When the psychological structure of other visual stimuli is understood, it will be possible to see if these sensory scanning strategies are general, or if they are restricted to behavior in response to natural language.

References

- Bever, T. G. Associations to S-R theories of language. In T. Dixon & D. Horton (Eds.), *Verbal behavior theory and its relations to S-R theory*. In press.
- Chomsky, N. *Aspects of the theory of syntax*. Cambridge: M.I.T. Press, 1965.
- Chomsky, N., & Halle, M. *The sound pattern of English*. New York: Harpers. Chap. 3. In press.
- Fodor, J. A., & Bever, T. G. The psychological reality of linguistic segments. *J. verbal Learn. verbal Behav.*, 1965, 4, 414-420.
- Garrett, M. The perceptual segmentation of speech. In T. Bever & W. Weksel (Eds.), *The structure and psychology of language*. New York: Holt, Rinehart & Winston. In press.
- Mackworth, N. A stand-mounted eye-marker camera for line of sight recording. *Percept. & Psychophys.*, 1967, 2, 119-127.
- Mehler, J., & Savin, H. Memory in the language user. In T. Bever & W. Weksel (Eds.), *The structure and psychology of language*. New York: Holt, Rinehart & Winston. In press.
- Miller, G. A. Some psychological studies of grammar. *Amer. Psychologist*, 1962, 17, 748-762.

Notes

- J. Mehler is a Lecturer and Research Associate at M.I.T., T. G. Bever is a Junior Fellow at Harvard and a Research Associate at M.I.T. and P. Carey is in the Psychology Department at Harvard. We thank Dr. N. Mackworth and the staff of the Harvard Center for Cognitive Studies for providing equipment, subjects, and advice. This research was supported by the National Institute of Mental Health No. MN-05120 to Harvard University, Center for Cognitive Studies and NIH: 5-TO1-HD00111-02 to Professor Morris Halle at M.I.T. We wish to express our gratitude to A. Battro for his helpful comments and dedication when this experiment was in the planning stage.
- For a review of this literature and the methodology see N. Mackworth, 1967.
- The one error in the balancing for letter-length is that "interesting" in DS+SS-2 is one letter longer than "surprising" in DS+SS-1. This error in stimulus preparation is taken into account in all phrases of this report. Each phrase was put in a longer sentence e.g., the actual stimulus sentence was "They gave her dog candies

with milk." See Note 7 for a listing of all the complete stimulus sentences. Notice that there is some difference in the underlying structures of the two interpretations of SS-1, 2 as well as the surface structure differences. Sentences SS-1, 2 are classified as having primarily surface structure differences because the underlying structure subject/object relations with the verbs are constant to either interpretation (i.e. in SS-1 "they told . . . stories" is common to either interpretation).

4. There is one suggestive asymmetry in the DS sentences. Structures DS-e, in which "of the hunters" is the deep structure subject of the verb, receive consistently more fixations than DS-f when "of the hunters" is the object. The effect is not large (a total of +6.3 on the last three words in favor of the subject interpretation) but it may be important. Other reading studies (T. G. Bower, personal communication) indicate that the deep structure subject is often encoded first in reading, regardless of its surface structure position.

5. We assume here that the "er" and "s" are not effective constituents themselves, and that the plural marker "s" does not add another phrase structure level. If it did, however, it would affect both kinds of words, equally, so the relative differences would remain.

6. The final syllable of "stories" and "farmer" is not included, since later constituents in the sentence affect them. The predicted value for the first syllable of "farmers" and "hunters" is 1.5, since it is an average of "farmers" with "authors" and "hunters" with "sisters"; see above. In SS-DS sentences, for purposes of the fixation pattern predictions the words "interesting" and "surprising" are assumed to have the structures ((interest)ing) ((surpris)ing).

7. The most notable points of failure to predict the direction correctly are at the ends of the critical sequences. This also may be due to the presence of words after the stimulus sequences. For the record, we present the stimuli as they appeared to the subject.

then they gave her dog candies with milk
then they told her cat stories with film
they are interesting farmers from Iowa
they are surprising authors from Peru
the shooting of the hunters was awful
the punching of the sisters was silly

The other prediction failures are in SS-DS-(d). This may be due to the length of the constituents, which would tend to blur any differences. (Note that the corresponding segments "erest" and "pris" in SS-DS sentences combine to count as 4½ letters from the fixation principle since one is five letters long and the other is four letters long. See Note 3.) (Values for single spaces between words are not included in the observed ratios since the data was insufficient to test them. In general, there were more fixations on spaces than this model predicts.)

8. A Pearson r was justified since the predictions and the results differed along a continuous scale.

9. The cyclic analysis of these eye-fixation patterns and of other similar psycholinguistic phenomena is introduced and discussed in T. G. Bever (in press). For the role of phrase structure in the linguistic analysis of English stress, see Chomsky and Halle (in press).