Journal of Experimental Psychology

Vol. 79, No. 3, Part 1 — March 1969

TRANSITIONAL PROBABILITY IS NOT A GENERAL MECHANISM FOR THE SEGMENTATION OF SPEECH[1]

T. G. BEVER,[2] J. R. LACKNER
Massachusetts Institute of Technology

AND W. STOLZ
University of Texas at Austin

The generality of a transitional probability model for the perceptual segmentation of speech was tested. Previous research has shown that clicks in sentences are subjectively located at boundaries between clauses. This might be a reflection of the low transitional probability between clauses rather than a demonstration that syntactic structure is actively used to organize speech processing. In the present experiment *S*s indicated the subjective location of a click presented during sentences which varied transitional probability within clauses. The results show that a click in the first word of a highly redundant (high-probability) two-word sequence tends to be subjectively perceived as occurring in the middle of the sequence, while a click in the second word of a low-probability two-word sequence tends to be perceived as occurring after the sequence. Thus high-probability sequences within clauses subjectively *attract* clicks into themselves, while low-probability sequences do not, indicating that transitional probability has different effects within and between clauses and thus is not a general mechanism for the active segmentation of speech.

Recent investigations of speech perception have shown that phrase structure is actively used to impose perceptual structure on the actual acoustic stimulus. For example, traditional and current linguistic theories agree that Sentence *a* is made up of two "clauses" (indicated by the parentheses).

(*a*) (John ran quite fast) (but we caught him in the narrow alley.)

A series of studies have shown that errors in the location of single clicks presented during sentences tend to be toward the point between clauses (Fodor & Bever, 1965; Garrett, Bever, & Fodor, 1966). For example, a click placed in "fast" or in "but" in the sentence above is located by a majority of *S*s as actually having occurred *between* those two words.

Several experiments have shown that this effect is not due to any actual pauses or intonational cues in the pronunciation of the sequences. First, Garrett, Bever, and Fodor (1966) used materials in which the identical

[1] This work was supported by Grant A. F. 19(628)-5705 and Grant SD-187 from the Department of Defense Advanced Research Projects Agency awarded to Harvard University. The authors are grateful to Harris Hartz for computation assistance and to H. L. Teuber for advice on the manuscript.

[2] Requests for reprints should be sent to Thomas G. Bever who is now at the Rockefeller University, Department of Psychology, New York, New York 10021.

acoustic sequence was assigned different clause structures depending on what preceded.[3] The results showed that the clause structure assigned each sequence "attracted" the subjective location of the clicks. In a second study, Abrams, Bever, and Garrett[4] found similar results with speech materials constructed by splicing words tape-recorded onto a random list.

The previous studies concluded that systematic errors in click location were due to the active use of syntactic structure to segregate the perceptual units of speech. An alternative view of speech processes would invoke sensitivity to transitional probability between words as the basis for click displacement. In the descriptive system using transitional probability, a word has a certain probability of occurrence given a string of preceding words. Any model of language which takes into account only the probability of a word or phrase (given any number of preceding words or phrases) is inadequate to describe language structure in general (cf. Chomsky & Miller, 1965). However, some psychological theories have attempted to use such a probabilistic mechanism to account for actual speech behavior. On this view, the "syntactic structuring" of speech is a structural derivative of peaks and troughs in interword transitional probability. For example, the transitional probability within clauses (e.g., between "quite" and "fast" in *a*) is characteristically higher than the transitional probability between a word ending one clause and the word beginning the next clause (e.g., between "fast" and "but" in *a*). To explain syntactic structure in this way is attractive to many psychologists since the notion of "transitional probability" is integral to Hullian and neo-Hullian theory and is used frequently in the description of other types of behavior. There have been several attempts to import this theoretical terminology into the description of language behavior (e.g., Johnson, 1965; Osgood, 1963; Osgood & Sebeok, 1954).

According to this sort of theory, the behavioral basis for the ongoing perceptual organization of speech is sensitivity to transitional probability. "Perceptual units" in speech are defined as sequences with high transitional probability bounded by points of low transitional probability. The point between clauses appears to have an independent psychological effect only because clause breaks in general coincide with points of low transitional probability. Hence the previous results that clause breaks subjectively attract interrupting clicks might also be taken as evidence for a probabilistic model of speech behavior. To test this critically, the authors examined the subjective location of clicks presented during high- and low-probability sequences with a constant syntactic structure. Results indicate that it is not *generally* the case that high-probability sequences repel subjectively interrupted clicks. In fact, if syntactic structure is held constant, the relative effect of high transitional probability sequences within clauses is subjectively to *attract* clicks from the extremes of the sequence toward the middle of the sequence.

Method

Materials.—Twelve sets of sentences were constructed. Each set contained two sentences with a point of high transitional probability and two sentences with a relatively low transitional probability at the same position in the sentence. The high and low transitional probability points in the sentences were generated by pairing two probable sequences of words and using their "cross-product" as the improbable sequences. For example, the cross-product of the probable sequences "men run" and "birds fly" yields the less probable sequences, "men fly" and "birds run." Sentential contexts were devised such that all four versions of each set could be inserted in the same pattern. For example, the four sequences appeared in the following set of four sentences: (*a*) Despite the high degree of stamina required some *men run* great distances. (*b*) Despite the high degree of stamina required, some *birds fly* great distances. (*c*) Despite the high degree of stamina required, some *men fly* great distances. (*d*) Despite the high degree of stamina required, some *birds run* great distances. Sentences *a* and *b* show the high-prob-

[3] Consider the sequence "eagerness to win the horse is quite immature." If it is preceded by "your . . ." then the clause break immediately follows "horse." But if that sequence is preceded by "In its . . ." then the clause break immediately follows "win." The authors cross-recorded one initial sequence or the other and tested *S*s on their ability to locate clicks in the different sentences.

[4] K. Abrams, T. G. Bever, & M. Garrett. Syntactic structure modifies attention during speech perception and recognition. Unpublished manuscript, 1968.

TABLE 1

SENTENCES USED IN THE EXPERIMENT

1.	a.	Often a wolf (howls) for many hours when he is alone.
	b.	baby (cries)
	c.	wolf (cries)
	d.	baby (howls)
2.	a.	To our surprise the lecturer (spoke) movingly and every one cried.
	b.	musician (played)
	c.	lecturer (played)
	d.	musician (spoke)
3.	a.	It is customary to change (clothes) before beginning to play.
	b.	shuffle (cards)
	c.	change (cards)
	d.	shuffle (clothes)
4.	a.	The talented author finished a hearty (meal) just a few hours ago.
	b.	thrilling (book)
	c.	hearty (book)
	d.	thrilling (meal)
5.	a.	You must change (trains) in order to travel to New York.
	b.	buy (tickets)
	c.	change (tickets)
	d.	buy (trains)
6.	a.	Despite the high degree of stamina required some birds (fly) great distances without resting.
	b.	men (run)
	c.	birds (run)
	d.	men (fly)
7.	a.	The prize winning dog wore a blue (ribbon) around its neck.
	b.	black (collar)
	c.	blue (collar)
	d.	black (ribbon)
8.	a.	While in Asia, the young man often saw chefs (cook) food and then give it away.
	b.	thieves (steal)
	c.	chefs (steal)
	d.	thieves (cook)
9.	a.	The children played up on the sandy (beach) most of the afternoon.
	b.	rocky (hill)
	c.	sandy (hill)
	d.	rocky (beach)
10.	a.	They found that the fire (burns) very fast when certain chemicals are sprayed on.
	b.	grass (grows)
	c.	fire (grows)
	d.	grass (burns)
11.	a.	Many people all over the world eat (rice) every day of their adult lives.
	b.	chew (gum)
	c.	eat (gum)
	d.	chew (rice)
12.	a.	She used to read (stories) to her children every night.
	b.	write (letters)
	c.	read letters)
	d.	write (stories)

Note.—Versions a and b are high transitional probability, c and d are low transitional probability.

ability sequences, Sentences *c* and *d* show the low-probability sequences.

The relative transitional probabilities of the sequences were tested in the sentence contexts using a sentence-completion procedure to insure that the high- and low-probability sequences were correctly designated (cf. Taylor, 1954). The sentence completions were run in two ways—using only the context preceding the second word of each critical sequence and using the context preceding and following the second word of the critical sequence. In each study *S*s responded in either a forced-choice or a free-choice mode. In the forced-choice mode, *S* chose which of the two alternatives was more appropriate (e.g., "run" or "fly" in the example). In the free-choice mode, *S* supplied any word which he thought was most appropriate in the second word position of each sequence. Thus, four tasks were used altogether; whole sentence, forced choice; whole sentence, free choice; left context, forced choice; and left context, free choice. The number of *S*s in each condition were 56, 60, 50, and 52, respectively. The results are summarized in Tables 1 and 2 with the sentences. The syntax

TABLE 2

PERCENTAGE OF *S*S CHOOSING THE CORRECT WORD

Sentence	Forced choice		Free choice	
	Left context	Left and right context	Left context	Left and right context
1. a.	100	89	44	71
b.	100	100	81	93
c.	0	11	0	16
d.	0	0	0	0
2. a.	40	100	4	69
b.	20	96	20	64
c.	60	0	0	0
d.	80	4	0	7
3. a.	72	87	63	83
b.	100	100	77	90
c.	28	13	0	0
d.	0	0	0	0
4. a.	96	70	59	40
b.	96	100	0	0
c.	4	30	0	0
d.	4	0	0	0
5. a.	92	93	0	0
b.	96	100	0	0
c.	8	7	0	0
d.	4	0	0	0
6. a.	32	100	27	93
b.	88	100	0	72
c.	68	0	0	0
d.	12	0	0	0
7. a.	88	79	69	83
b.	96	93	70	90
c.	12	21	19	17
d.	4	7	15	10
8. a.	80	77	56	44
b.	92	100	35	86
c.	20	23	0	0
d.	8	0	0	0
9. a.	32	62	48	48
b.	88	89	27	40
c.	68	38	30	38
d.	12	11	0	0
10. a.	72	100	27	40
b.	80	97	42	77
c.	28	0	0	6
d.	20	3	0	0
11. a.	96	97	19	30
b.	96	93	86	62
c.	4	3	0	0
d.	4	7	0	0
12. a.	96	100	0	72
b.	48	100	27	86
c.	4	0	0	0
d.	52	0	0	0

Note.—Correct word is parenthesized in Table 1.

TABLE 3

EXPERIMENTAL DESIGN WITH EXAMPLE SEQUENCES

Location of click	High transitional probability		Low transitional probability	
	Spliced	Nonspliced	Spliced	Nonspliced
First word of sequence	...mén	run....	...mén	fly....
	...biŕds	fly....	...biŕds	run....
Second word of sequence	...men	rún....	...men	flý....
	...birds	flý....	...birds	rún....

Note.—′ indicates the objective placement of the click.

was held constant within each set of four sentences. Five sets involved a verb-noun sequence (e.g., "write letters"), four a noun-verb sequence (e.g., "birds fly"), and three an adjective-noun sequence (e.g., "hearty meal"). The materials were roughly balanced for left-right position of the sequence in the sentence.

Each of the 48 experimental sentences (12 different sets of the 4 shown previously) was initially recorded with subdued intonation and then copies were recorded onto one track of a stereophonic tape. Each of the 48 sentences had 4 experimental versions. In two versions a click was located on the blank track of the recording, coincident with the first word of the critical two-word sequences ("men" and "birds" in six). The other two copies of each sentence had a click coincident with the second word of the critical two-word sequence ("run" and "fly" in six).

For each click position, sentences were recorded in two ways: (*a*) normally or (*b*) by splicing the first half of one version of a sentence onto the second half of another version of the sentence. For example, if the italicized portions of *"men run great distances"* and *"birds fly great distances"* are exchanged by tape-splicing, then copies of the two relatively low-probability sentences result. To control for a possible effect of the splice itself, a splice was introduced into a copy of each of the high-probability versions as well. Thus the spliced sentences controlled for any possible systematic differences in pronunciation between probable and improbable sequences.

These procedures yielded 16 experimental sentences for each set of four basic sentences. (Table 3 presents the matrix of the independent variables.) Sixteen experimental lists were constructed, each containing one of the experimental sentences from each of the 12 sets of 16. Each of the 12 sets of four was represented in the same order (that of Tables 1 and 2) in every experimental list (although in each list by a different experimental version). The lists were balanced for high and low transitional probability of the critical sequence, click position, and method of preparation.

Each experimental list included an additional six sentences with different syntactic structures. This was to insure that our results would be comparable to those of previous studies on the effect of syntactic structure on click location. In the present experiment there were six pairs of sentences which differed critically in their syntactic structure. The sentences contained different kinds of verb complementation, as in the following examples:

[They will *defy John*] [to hit the ball]—
"verb-phrase complements,"
[They will *desire*] [*John* to hit the ball]—
"noun-phrase complements."

Transformational linguistic analysis assigns these structures different syntactic segmentation (represented by the brackets) despite their superficial similarity (Rosenbaum, 1967). Clicks should be located subjectively between the two italicized words more often for the second sentence than for the first. Six pairs of sentences similar in structure were constructed (see Table 4). Clicks were placed either in the verb or the following noun. For each click position, experimental versions were produced either by normal reading, or tape splicing. Each of the 16 experimental lists had examples of both kinds of complement structure, randomly assigned from the six pairs.

Experimental procedure.—The click was about 25 msec. in duration and equal to the loudest speech sound. The sentence was played into one Jenson HS-1 earphone and the click into the other. The experimental procedure and instructions were exactly those used in Fodor and Bever (1964). Two hundred and fifty *S*s (run in groups of five) were instructed to listen to each sentence, write it down following the auditory presentation, and indicate with a slash, "/," where in the sentence the single click had occurred. For each experimental order, half of the *S*s heard the speech in the right ear and the click in the left, and half had the reverse earphone orientation. The *S*s were right-handed Massachusetts Institute of Technology undergraduates, native speakers of English, who volunteered for paid participation.

Results

Sixty-eight percent of the click location responses to the 12 sentences which varied transitional probability were incorrect for sentences with high-probability sequences and 69% for the sentences with low-probability sequences. To test the effect of transitional probability on click location, location errors moving the click subjectively into the position between the two critical words were compared with errors of equal magnitude moving the click out of the sequences. For example, if a click objectively in "fly" was reported as occurring between "fly" and "men," then this error was counted as "attracted" into the sequence. If the click was reported as occurring between "some" and "men" the error was counted as "repelled" out of the sequence.[5]

This analysis shows that high-probability sequences attract click locations rather than repel them. Seventy percent of the errors in response to the high-probability versions were attracted into the sequence and only 49% for the improbable versions. The strength of this result was tested by comparing the two probable versions (e.g., "men run . . ." and "birds fly . . .") for each of the 12 sets of four sentences. Of these 12

[5] Notice that if the critical sequence contained a polysyllabic word (e.g., as in "shuffle cards") then the only errors in locating the click in "shuffle" bearing on the relative ratio are errors of ± 1 syllable rather than $\pm \frac{1}{2}$ syllable.

For a fuller discussion of the results in response to the sentences in Table 4 presented in this paper see Bever, Lackner, & Kirk (in press).

TABLE 4

Sentence Pairs with Different Syntactic Structures

1. a. The general defied the troops to fight against the advancing enemy.
   b. preferred
2. a. The little girl told her mother to cook her some cauliflower.
   b. hated
3. a. The shopkeeper directed John to pile some boxes in the corner.
   b. desired
4. a. The teacher tempted the guilty boy to inform on his classmates.
   wanted
5. a. The corrupt police can't force criminals to confess very quickly.
   b. bear
6. a. The prophet will cause the people to renounce their indifference.
   b. like

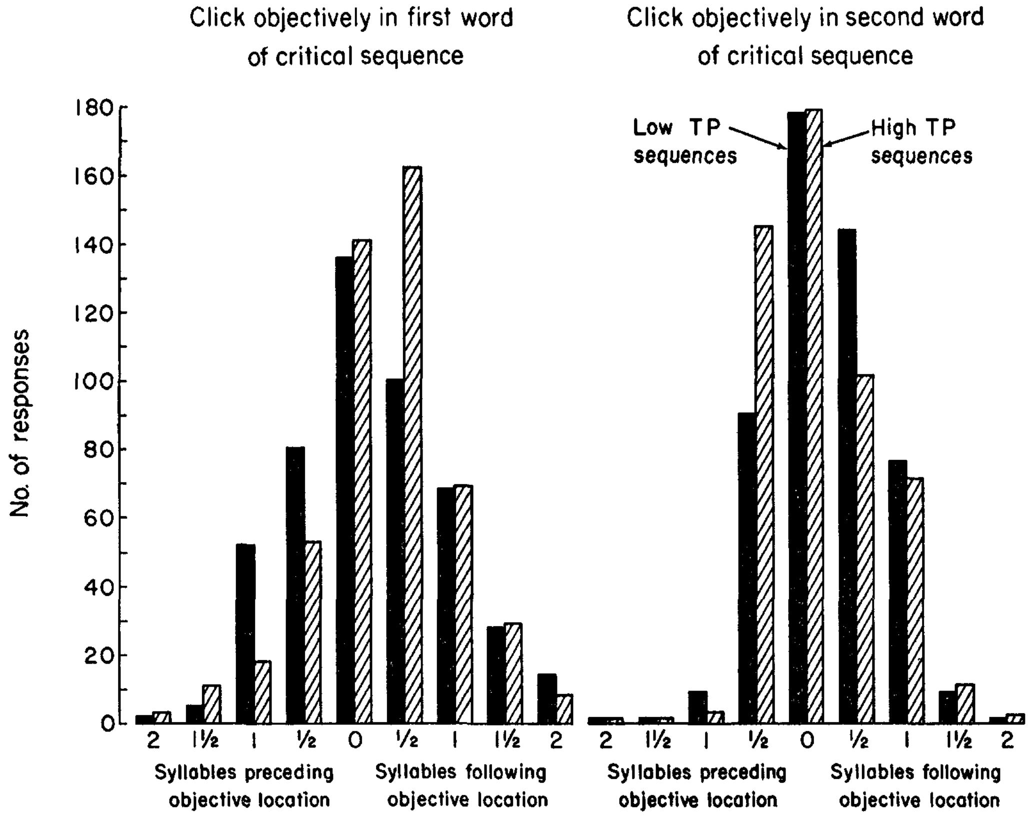

FIG. 1. Subjective location of clicks in objectively high- and low-probability sequences. (Solid bars indicate number of responses to low transitional probability sequences, striped bars indicate number of responses to high transitional probability sequences.)

comparisons all showed a higher proportion of errors into the sequence for the relatively probable versions than for the improbable versions ($p < .001$, sign test, two-tailed). This result was as true for the "spliced" ($p < .001$) as for "nonspliced" ($p < .001$) versions of the sentences.[6] Furthermore, this result was significant by *S*. Two *S*s (one who heard the speech in the left ear, one with the reverse orientation) were selected randomly from each of the 16 experimental groups. Of these a significant majority gave more errors into the sequence for the probable versions ($p < .01$, two-tailed sign test by *S*).

The fact that relatively more sequence-intrusion errors occur for probable sequences is due to two biases in the subjective location of clicks: (*a*) clicks objectively in the first word of probable versions are attracted into the sequence and (*b*) clicks objectively in the second word of improbable versions are repelled out of the sequence. Figure 1 presents the distribution of responses for each of the four click-location probability combinations. The distributions for the probable version with the click in the first word and the improbable version with the click in the second word are skewed to the left of the actual click position while the responses to the other two conditions are not skewed. These asymmetries are significant by sentence ($p < .01$, sign test, two-tailed, for each comparison in the 24 probable- and 24 improbable-sequence sentences). These asymmetries are also significant by *S*: 27 *S*s of the 32 randomly selected mislocated more

[6] There were no overall differences in response to the "spliced" and "normal" versions so this distinction was dropped in all further analyses.

clicks to the right than left for the clicks objectively in the first word of probable sequences while only 3 showed the reverse ($p < .01$, two-tailed); 19 *S*s mislocated more clicks to the right for clicks objectively in the second word of improbable sequences while only 9 showed the reverse ($p < .10$, two-tailed). The responses to the clicks objectively in the first word of an improbable sequence and in the second word of a probable sequence are not skewed in either direction.

The errors in location of clicks in the syntactically differentiated sentences replicate the earlier findings that syntactic structure affects the subjective location of clicks. In this experiment the syntactic structure made one prediction: there should be relatively more errors placing the click between the verb and the following word for noun-phrase complement sentences (with verbs like "desire") than for verb-phrase complement sentences (with verbs like "defy"). This was tested by computing for each unique sentence–click combination the number of location errors attracted into that position and the number of errors of equal distance away from that position (see Footnote 5). The proportion of attraction errors was higher for the noun-phrase complement (.8) than for the verb-phrase complement (.6). Although small, this difference was maintained for each of the six sentence pairs ($p < .02$, one-tail, by sign test). Furthermore, there were significantly more *S*s (of the 32 randomly chosen ones) who had higher proportions of errors into the verb–noun sequence for the noun-phrase sentences they heard than for the verb-phrase sentences ($p < .01$, sign test, two-tailed).

The probable and improbable sequences involved three different kinds of within-clause syntactic phrase structures: subject–verb, verb–object, and adjective–noun (see Tables 1 and 2). These different phrase structures involve within-phrase phrase-structure boundaries of three relative "depths" corresponding to the analysis of this example:

(the nice boy) (hit (the girl))
0 2 1

Adjective–noun phrase-structure breaks are the least major breaks, (no phrase divisions); verb–object breaks are larger (one division); and subject–verb breaks are the largest (two divisions). If all levels of phrase-structure divisions affect click placement, then there should be differences in the proportion of errors into a sequence depending on the phrase structure of the sequence. However, there was no difference in the tendency for attraction errors for these different syntactic structures independent of transitional probability. This is evidence that "minor," within-clause phrase-structure breaks do not affect click location (see Bever, Lackner, & Kirk, in press, for other data confirming this).

As in our previous experiments on the location of clicks in speech, there was an asymmetry correlated with the ear orientation of the speech and click. In general, a click heard in the left ear was located subjectively at a point preceding the subjective location of a click heard in the right ear. For the 12 sets of sentences which varied probability, the mean location for clicks heard in the right ear (with speech in the left) was .37 syllables later than the actual click location; for clicks in the left ear (with speech in the right), .12 syllables later than the actual location (significance of difference, $p < .01$, sign test, two-tailed, by sentence).

There was a slight asymmetry in the effect of high- and low-probability sequences associated with earphone orientation. For those selected *S*s who heard the click in the left ear, the mean percentage of attraction errors for high-probability sequences was 71% and for low-probability sequences, 45%. This yielded a difference of 26% associated with the difference in sequence probability for *S*s who heard the click in the left ear and the speech in the right ear. In contrast, *S*s who heard the click in the right ear and the speech in the left ear had 69% attraction errors for high-probability sequences and 54% attraction errors for low-probability sequences, to give a smaller difference of 15%. Thus the effect of sequence probability was greater for clicks in the left ear than in the reverse orientation ($p < .04$, two-tailed, by sentence).

Discussion

These results show that transitional probability has a different effect on click location within clauses than the previously found effect between clauses. Previous experiments have shown that the low transitional probability which occurs at the point between two clauses attracts the subjective location of clicks; the present experiment shows that, within clauses, relatively high transitional probability attracts the subjective location of clicks. The main finding that high-probability sequences attract clicks is contradictory with the view that psychological units are defined as highly associated sequences: points of high transitional probability could not correspond to the points *between* processing units.

These findings support a different view of speech perception and its interaction with click location. Bever (in press) has suggested that listening to speech involves two competing activities—attending to the external acoustic stimulus and developing an internal perceptual representation of it. Abrams, Bever, and Garrett (see Footnote 4) suggest further that errors in click location cluster just at those points when the internal perceptual analysis of speech occurs. The most obvious point for such a perceptual analysis to occur is between clauses, since the clause represents a basic unit of speech.

If it is generally true that clicks tend to be perceptually displaced toward points where perceptual analysis of speech occurs, then the present experiment indicates that perceptual proccessing tends to occur at points of high predictability within clauses as well as occurring between clauses. Thus, when a listener already has strong expectations about what he will hear next, he may only superficially monitor the incoming signal and simultaneously process the developing clause (and the click). In the present materials, the first word of the experimental two-word sequence served to establish a strong expectation, while the second one either confirmed it (high probability) or disconfirmed it (low probability). In the former case, attention to the second expected word could be low, allowing accurate perception of a click in that word or a systematic delay in perceiving a click in the first word. Exactly these phenomena were observed. During a low-predictable sequence, a click in the first word should not tend to shift perceptually to the second word, since the second word should require a relatively large amount of attention (to counteract the strong expectancy built up by the first word for another word). Similarly, a click occurring objectively in the second word of a low-predictable sequence may be expected to be displaced perceptually toward a part of the utterance which is more predictable (and thus which requires less complete monitoring). The present results also conformed to this hypothesis, with no bias observed in the perception of a click in the first word and a delay observed in the perceptual location of a click in the second word in low-predictable sequences.

Thus it is clear that listeners actively use knowledge *both* of syntactic structure and of transitional probabilities during speech perception. However, transitional probability has one sort of effect on perceptual grouping between clauses and the opposite effect within clauses. Hence, transitional probability does not provide a general account for the behavioral segmentation of speech.

REFERENCES

Bever, T. G. The cognitive basis for linguistic universals. In J. R. Hayes (Ed.), *Cognition and the development of language.* Englewood Cliffs, N.J.: Prentice Hall, in press.

Bever, T. G., Lackner, J., & Kirk, R. The underlying structures of sentences are the primary units of speech processing. *Perception and Psychophysics,* in press.

Chomsky, N., & Miller, G. A. Introduction to the formal analysis of natural languages. *Handbook of mathematical psychology.* New York: Wiley, 1965.

Fodor, J. A., & Bever, T. G. The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior,* 1965, **4**, 414–420.

Garrett, M., Bever, T. G., & Fodor, J. A. The active use of grammar in speech perception. *Perception and Psychophysics,* 1966, **1**, 30–32.

Johnson, N. F. The psychological reality of phrase-structure rules. *Journal of Verbal Learning and Verbal Behavior,* 1965, **4**, 469–475.

Osgood, C. E. On understanding and creating sentences. *American Psychologist,* 1963, **18**, 735–751.

Osgood, C. E., & Sebeok, T. A. (Eds.), *Psycholinguistics.* Bloomington: Indiana University Press, 1954.

Rosenbaum, P. *The predicate complement construction of English.* Cambridge: M.I.T. Press, 1967.

Taylor, W. L. Application of "cloze" and entropy measures to the study of contextual constraint in samples of continuous prose. Unpublished doctoral dissertation, University of Illinois, 1954.

(Received April 17, 1968)