

Opposed Effects of a Delayed Channel on Perception of Dichotically and Monotically Presented CV Syllables*

M. Studdert-Kennedy,⁺ D. Shankweiler,⁺⁺ and S. Schulman
Haskins Laboratories, New Haven

We wish to report a new phenomenon in binaural speech listening that we have termed the "lag effect." The effect is seen in the greater accuracy with which subjects identify the lagging member of a pair of temporally overlapped syllables presented to opposite ears. Earlier experiments had shown that if CV or CVC syllables, differing only in their initial or final consonants, were presented in dichotic competition, those presented to the right ear were correctly reported significantly more often than those presented to the left (Shankweiler and Studdert-Kennedy, 1967; Studdert-Kennedy and Shankweiler, 1970). As part of a general program of research into the conditions of this right ear advantage for consonants, we undertook to titrate the effect in temporal units: our plan was to estimate the number of milliseconds by which the left ear syllable should lead the right ear syllable for the right ear advantage to be abolished. In the event, we found that the right ear advantage was more readily abolished by a left ear lag than by a left ear lead. The effect has now been repeatedly confirmed both at our own laboratory and elsewhere (Berlin et al., 1970; Lowe et al., 1970). Here we wish simply to report some of its conditions as uncovered in the original experiment.

The stimuli were formed from the syllables /ba, da, ga, pa, ta, ka/, synthesized on the Haskins Laboratories Parallel Formant Synthesizer and each 250 msec long. Syllables were recorded in pairs, one on each channel of a balanced two-track tape recorder. By means of a computer-aided routine, two 240-pair random order tapes were prepared: the onset of one member of each pair was made to lead (or lag) the onset of the other by 0, 5, 10, 20, 25, 50, 70, or 120 msec. The two tapes provided a fully balanced 480-item

*This paper appeared in *J. Acoust. Soc. Amer.* 48, 599-602 (1970).

⁺Also, Queens College, City University of New York, Flushing.

⁺⁺Also, University of Connecticut, Storrs.

test in which each syllable occurred equally often on each channel, paired with each syllable other than itself, for a total of thirty presentations at each lead and lag value other than zero, at which it occurred sixty times. These tapes were intended for dichotic presentation. A second pair of tapes was prepared for monotic presentation by mixing two channels of the dichotic tape electronically and recording the output on a single track.

The subjects were sixteen, right-handed, undergraduate women, all of whom had scored 95% or better on both ears in monaural identification tests of the synthetic syllables. As in previous dichotic studies (Shankweiler and Studdert-Kennedy, 1967; Studdert-Kennedy and Shankweiler, 1970), appropriate counterbalancing procedures distributed all effects due to recorder channels, earphone characteristics, positions of earphones on the head, or sequence of testing equally over the ears of the entire group of subjects. Subjects were instructed to record two from the set of six consonants on each trial, writing on an answer sheet and guessing if necessary.

As a baseline against which the dichotic data may be judged, we first present the group monotic data: in Figure 1, mean percent correct is plotted as a function of temporal lag (negative) and lead (positive) in milliseconds, for right and left ears. Each point is based on 480 judgments (960 at 0 msec). The two ears give essentially identical results: performance is at chance level for syllables with onsets that lag by 25 msec or more but then rises steadily to virtually perfect performance for syllables that lead by 50 msec or more. The functions were similar for all subjects: every one of the sixteen reached at least 95% correct for a lead of 50 msec.

The results seem open to a straightforward peripheral masking interpretation. Although each syllable was approximately 250 msec long, the important cues for the identification of its initial consonant occur in the first 50 msec, during which the syllable is also rising to its maximum amplitude. As lead time is increased from zero, more and more of the crucial portion of the syllable is presented without interference from the lagging syllable until, at 50 msec, all needed consonantal information in the leading syllable is freely available, and performance is almost perfect. On the other hand, as lag time is increased from zero, more and more of the crucial portion of the lagging syllable occurs during the period of maximum amplitude of the leading syllable until, at -50 msec, the important cues in the lagging syllable are fully masked and performance drops to chance. This account squares with the subjective impression of the monotic pairs at the longer lead/lag values: one hears a single

Mean percent correct by ear on monotically presented CV syllable pairs as a function of temporal lead in milliseconds. For sixteen subjects.

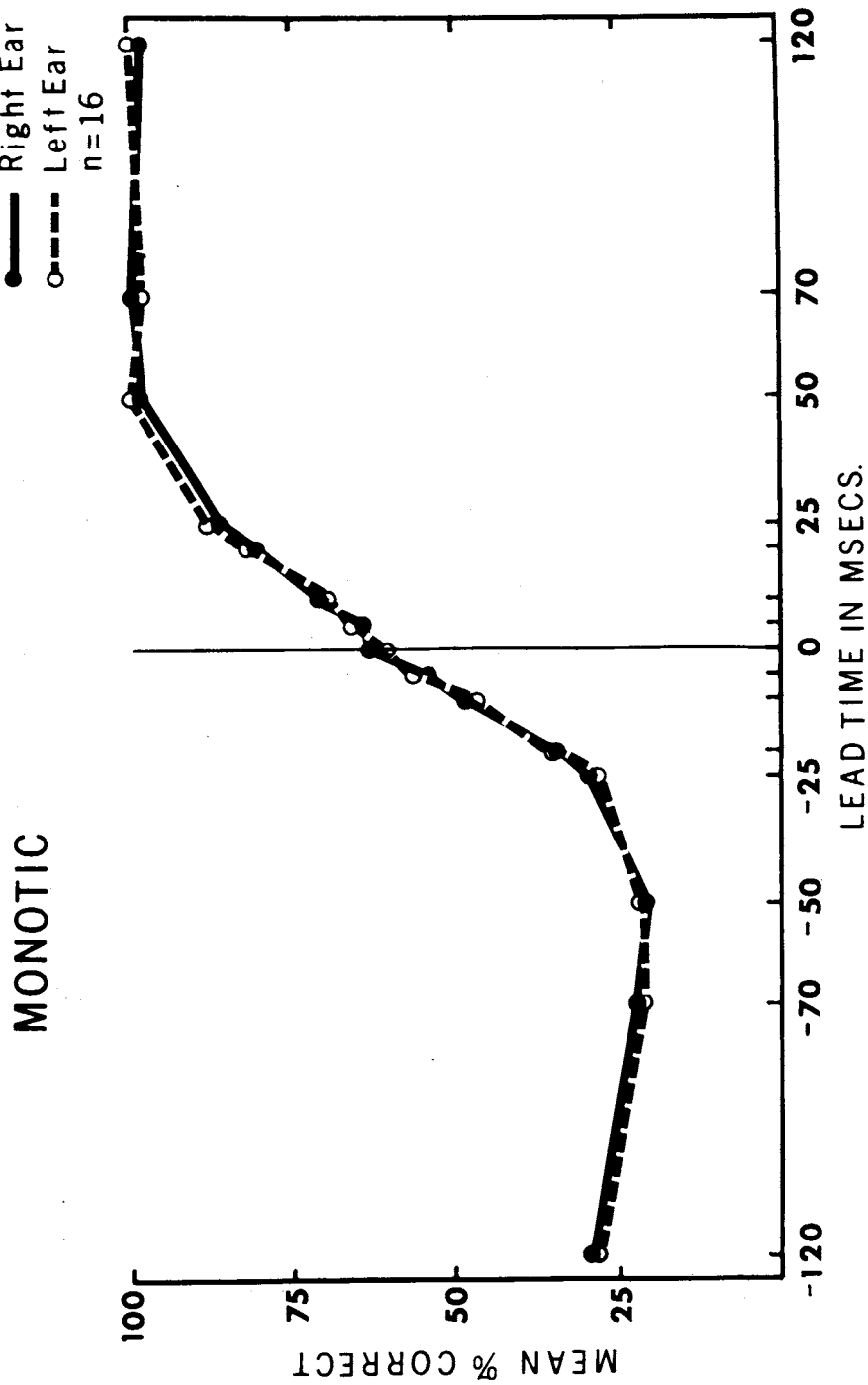


FIG. 1

Mean percent correct by ear on dichotically presented CV syllable pairs as a function of temporal lead in milliseconds. For sixteen subjects.

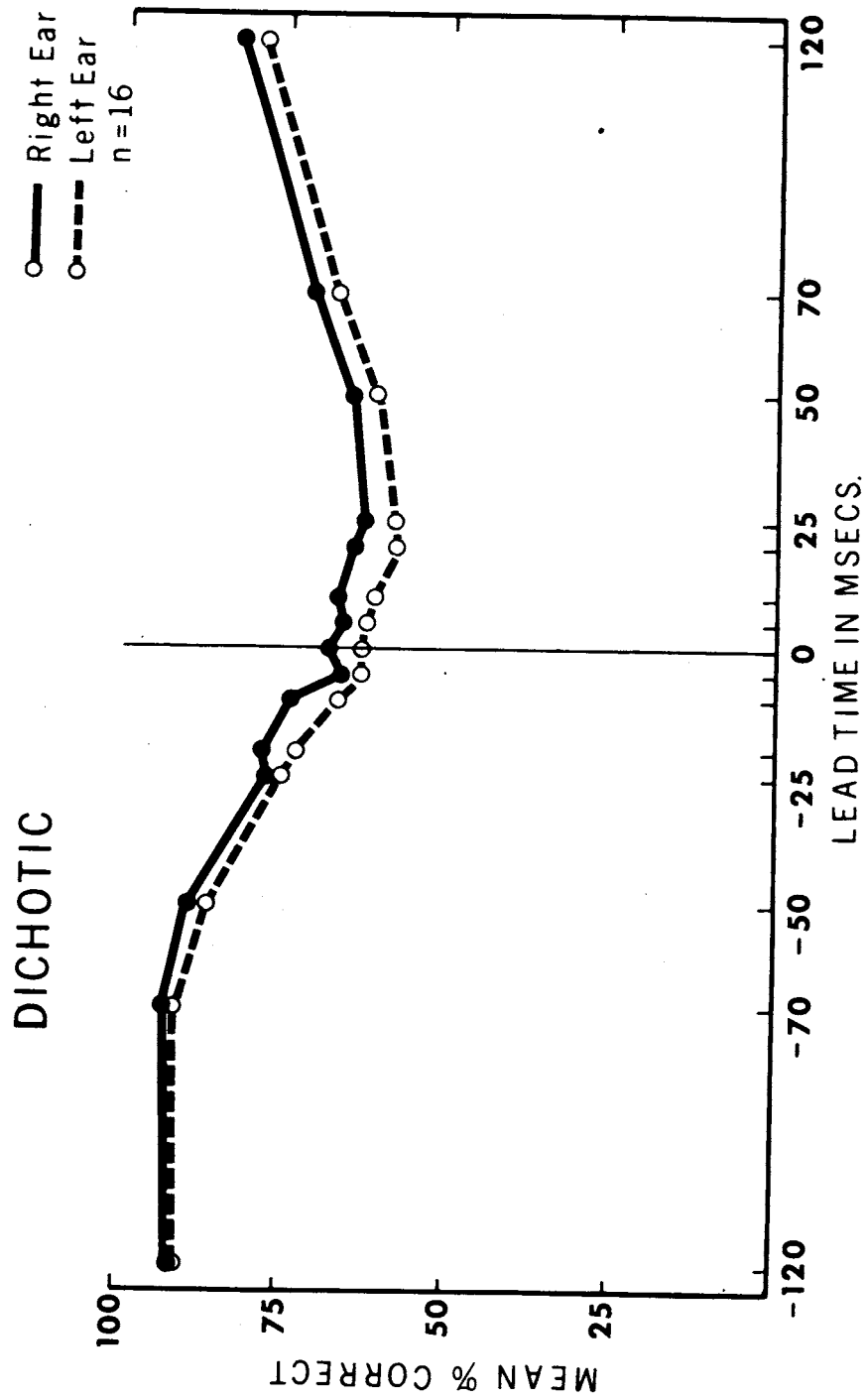


FIG. 2

syllable with a superimposed click.

The dichotic results present a quite different picture. Figure 2 displays the group dichotic results plotted on the same coordinates as Figure 1. On this plot, the difference between levels for left and right ears is a measure of the ear advantage (laterality effect), and the slopes of the functions from their minima measure the advantages accruing from changes in lead or lag time. Where the two functions are parallel, laterality effect and temporal effects are additive; significant deviations from the parallel indicate some interaction between the two effects.

We note first a clear laterality effect: right ear performance is superior to left at every lag/lead value other than -120 msec. Ten of the sixteen subjects show significant right ear advantages by matched pair t-tests over the lag/lead range; four show no significant ear advantage; two show significant left ear advantages. Subject by ear interaction is significant by analysis of variance, hence the overall ear effect is not significant. Individual differences of this order are common in dichotic experiments and may be related to differences in cerebral language dominance. Figure 3 gives some idea of the variability: examples of a clearly right-eared subject (above) and of a subject showing no significant ear advantage (below).

Second, we note that increases in the amount of lag yield, for both ears, increases rather than decreases in performance. Furthermore, the functions are not symmetrical: they reach their minima at lead values of 20 or 25 msec, rather than at zero; they reach their maxima at a lag value of -70 msec, where performance is superior by some 20% to performance at the corresponding lead value. In other words, the functions climb more rapidly over the lag than over the lead range. And this is true of every subject, despite considerably greater intersubject variability in the dichotic than in the monotic data. The overall effect of temporal offset is highly significant by analysis of variance, and there is no significant subject by temporal offset interaction.

The advantage of the lagging over the leading syllable may be more clearly seen if we replot the data of Figure 2 so that each pair of points shows the mean percent correct by ear for all trials of a given type. For example, the pair of points at the extreme left in Figure 4 gives performance for trials on which the left ear lagged by 120 msec, and the corresponding pair at the extreme right gives performance for trials on which the right ear lagged by 120 msec. (Figure 4 may be generated by rotating the right ear function of

Percent correct by ear for two subjects on dichotically presented CV syllable pairs as a function of temporal lead in milliseconds.

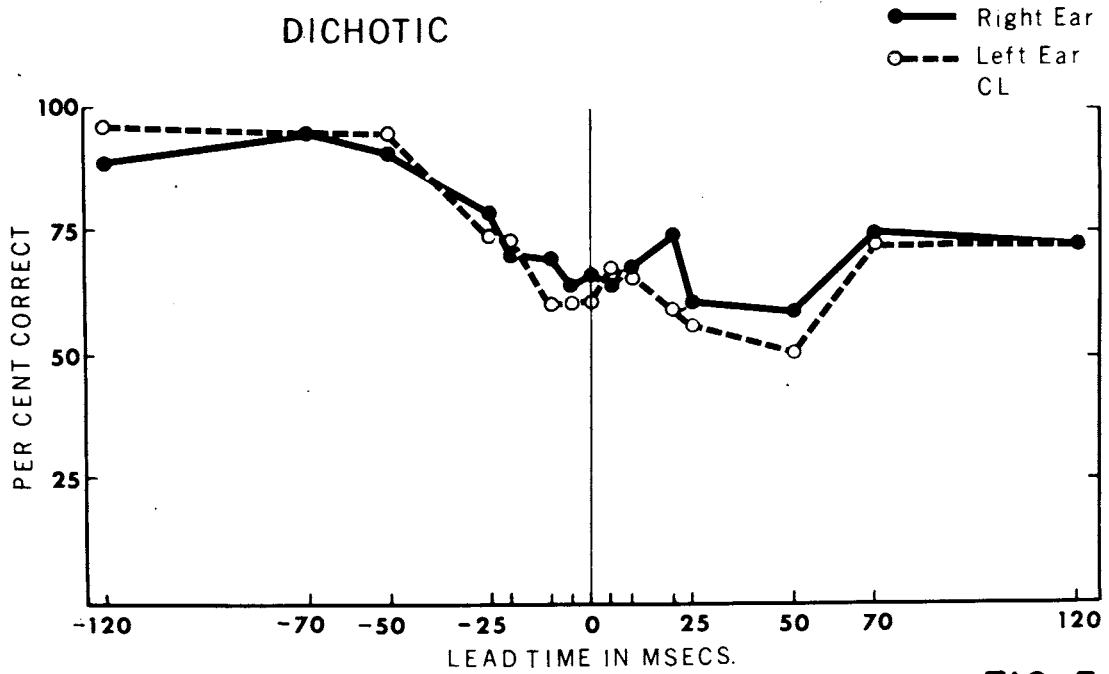
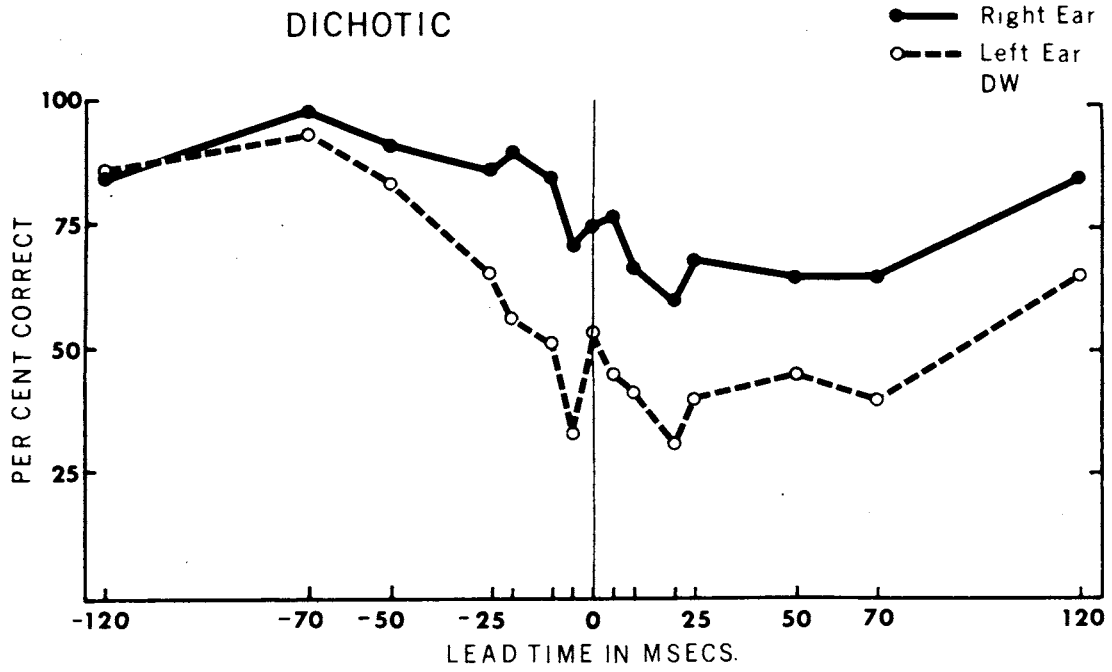


FIG. 3

Mean percent correct by ear on left lag and right lag CV syllable pairs, dichotically presented. For sixteen subjects.

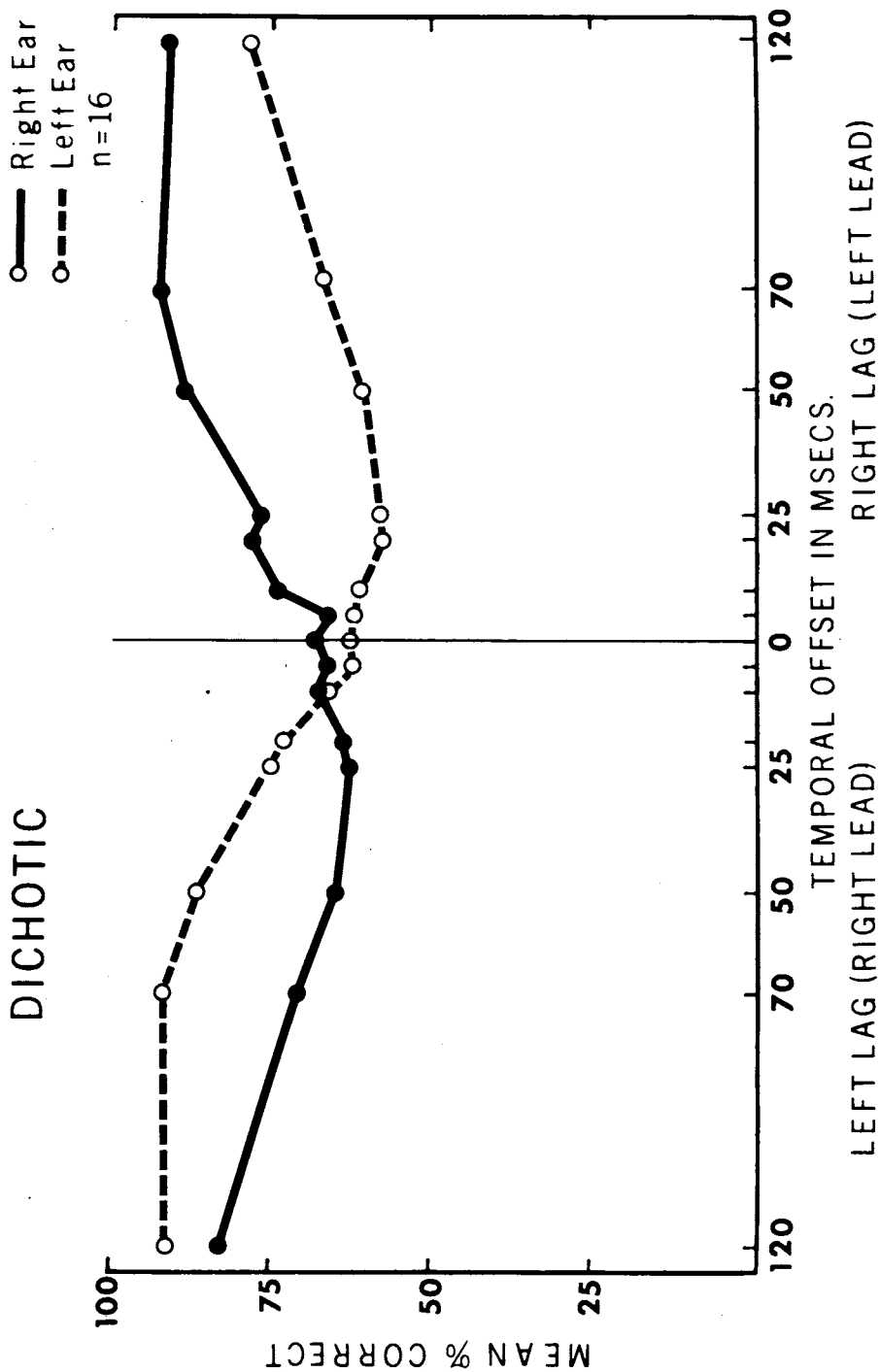


FIG. 4

Figure 2 through 180 degrees in a plane vertical to the page.) We see immediately that the ear to which the lagging syllable is presented almost invariably has the advantage over the leading ear. The exception is over the short, left ear lags (0-10 msec), where the right ear advantage under dichotic stimulation is sufficient to cancel the left ear advantage from lag. In fact, for these group data, 10 msec is the titration value that we originally sought, that is, the temporal advantage to the left ear necessary to cancel the dichotic advantage to the right. However, the value is not reliable across subjects.

Cancellation of the right ear advantage by an appropriate left ear lag suggests that the laterality and lag effects are independent phenomena. The same conclusion is suggested by the asymmetry of Figure 4: the wider separation of the two curves over the right lag range than over the left is due to the fact that the right ear, whether leading or lagging, has an overall higher level of performance than the left. The generally parallel courses of the two curves in Figure 2 makes the same point, and analysis of variance shows no significant interaction between ear and temporal offset.

We may now pose the problem raised by the dichotic lag effect fairly precisely. From Figure 2 it is evident that there is little variation in performance between -5 and +50 msec; within this range, the functions for both ears reach a broad minimum. For fifteen out of sixteen subjects this is the range within which both ears reach their minima; the sixteenth subject gives her minima at +70 msec. Thus, for every subject, dichotic performance is at its worst in the very range of lead values over which monotic performance is at, or rising to, its peak. The paradox sharpens when we recall that the conditions of presentation for the leading portion of the leading syllable are identical under monotic and dichotic presentation. For example, under both conditions, the initial 50 msec of a syllable leading by that amount are presented without interruption to a single ear. These 50 msec carry all the information needed for identification of the initial consonant, and under monotic conditions, virtually perfect identification is achieved by every subject, while performance on the syllable that lags by 50 msec drops to chance. Under dichotic conditions, on the other hand, performance on the leading syllable is, for every subject, close to her function minimum and on the lagging syllable close to her function maximum.

What gives rise to this reversal of the direction of the effect under dichotic conditions? The question is of interest for the light that its answer may throw on the processes of speech perception. For while the monotic lead

effect may be interpreted as an instance of peripheral, simultaneous masking, the dichotic lag effect seems to be of central origin, possibly analogous to metacontrast effects in vision. Werner (1935) showed that perception of a disc flashed on a screen might be blocked if rapidly followed by presentation of a ring having the same internal diameter as the disc. He attributed the effect to interference by the ring with development of the disc's contour. Later work (for example, Kolers and Rosner, 1960) showed that a similar effect might be obtained dichoptically and hence, that it involved central mechanisms.

Interpretation of the dichotic lag effect along analogous lines would assume processing of the important cues in the leading stimulus to be incomplete at the time the lagging stimulus arrived along a different channel to compete for, and frequently capture, the processors. Occlusion of the leading syllable by a switch in channels just as the crucial information in that syllable is being processed recalls the finding of Huggins (1964) that the rate of across-ears switching most disruptive to speech perception is roughly equal to the syllable rate. A similar disruption does not occur when the lagging syllable is presented along the same channel in wake of the first, presumably because it is masked at a peripheral point in the pathway.

The notion that the lag effect reflects interruption of speech processing is further suggested by control data. Studies with nonspeech have not yet been completed, but Porter, Shankweiler, and Liberman (1969) have reported that, if the stimuli are steady-state synthetic vowels, the advantage tends to the leading, rather than to the lagging, stimulus. Given that such stimuli have been found, under other experimental conditions, to be perceived in the manner more of nonspeech than of speech (Liberman et al., 1967; Shankweiler and Studdert-Kennedy, 1967; Studdert-Kennedy and Shankweiler, 1970; Studdert-Kennedy et al., 1970), we may reasonably suspect that the lag effect is tied to speech and, specifically, to those components of the speech stream for which a relatively complex decoding operation is necessary.

However, an adequate account of the effect and of its implications for speech perception calls for much further study. Several experiments are already under way at Haskins Laboratories. These include studies of individual differences, nonspeech controls, attention switching, channel tracking, and consonant feature errors.

References

- Berlin, C.I., Willett, M.E., Thompson, C., Cullen, J.K., and Lowe, S.S. (1970) Voiceless versus voiced CV perception in dichotic and monotic listening. *J. Acoust. Soc. Amer.* 47, 75 (A).
- Huggins, A.W.F. (1964) Distortion of the temporal pattern of speech: interruption and alteration. *J. Acoust. Soc. Amer.* 36, 1055-1064.
- Kolers, P. and Rosner, B.S. (1960) On visual masking (metacontrast): dichoptic observation. *Amer. J. Psychol.* 73, 2-21.
- Liberman, A.M., Cooper, F.S., Shankweiler, D., and Studdert-Kennedy, M. (1967) Perception of the speech code. *Psychol. Rev.* 74, 431-461.
- Lowe, S.S., Cullen, J.K., Thompson, C., Berlin, C.I., Kirkpatrick, L.L., and Ryan, J.T. (1970) Dichotic and monotic simultaneous and time-staggered speech. *J. Acoust. Soc. Amer.* 47, 76 (A).
- Porter, R., Shankweiler, D., and Liberman, A.M. (1969) Differential effects of binaural time differences in perception of stop consonants and vowels. Proc. 77th Annual Convention of the Amer. Psychol. Assn.
- Shankweiler, D. and Studdert-Kennedy, M. (1967) Identification of consonants and vowels presented to left and right ears. *Quart. J. Exptl. Psychol.* 19, 59-63.
- Studdert-Kennedy, M., Liberman, A.M., Harris, K.S., and Cooper, F.S. (1970) The motor theory of speech perception: a reply to Lane's critical review. *Psychol. Rev.* 77, 234-249.
- Studdert-Kennedy, M., and Shankweiler, D. (1970) Hemispheric specialization for speech perception. *J. Acoust. Soc. Amer.* 48, 579-594.
- Werner, H. (1935) Studies in contour: I. quantitative analyses. *Amer. J. Psychol.* 47, 40-64.