

## Stimulus Dominance and Ear Dominance in the Perception of Dichotic Voicing Contrasts

BRUNO H. REPP

*Haskins Laboratories*

Two studies were conducted to determine the effect of variations in voice onset time (VOT) on the perception of dichotic stop-consonant-vowel syllables contrasting in the voicing feature. The dichotic stimuli were partially fused, so that only a single response was required. Variations in VOT had a systematic effect on the probability of hearing the fused stimuli as voiced or voiceless. Changing the VOT of a voiceless stimulus had a larger effect than changing the VOT of a voiced stimulus. Unless one of the competing stimuli was close to the category boundary, the perceptual integration of their VOTs seemed to be roughly additive. The relative phase of the periodic portions of the stimuli had an unexpected effect on perception that remains to be explained. A number of subjects showed very strong right-ear dominance in these tests. The range and reliability of the laterality effects obtained, as well as certain other methodological features, make the present tests promising as tools for assessing individual differences in ear dominance.

### INTRODUCTION

When two different auditory stimuli are presented simultaneously to the two ears, the perceptual result depends on a number of factors. One of these is dichotic (binaural, lower-level) fusion. It determines whether one or two stimuli are heard. If the two inputs are very similar in their spectral and temporal characteristics, they may fuse, so that only a single stimulus is heard. If the two stimuli are dissimilar, two separate events are heard at the two ears, but it may nevertheless be difficult to identify both of them correctly. Two other factors determine which of the two competing inputs is perceptually more prominent. "Perceptual prominence" means, in the case of fused stimuli, that the fused percept sounds more like one component than the other or, in the case of unfused stimuli, that one of the inputs is more often correctly identified (or stands out more clearly) than the other. One of the two factors affecting perceptual

Experiment I was supported by NIH Grant DE00202 to University of Connecticut Health Center, Farmington, and was made possible by the hospitality of Haskins Laboratories. Experiment II and preparation of this paper were supported by NICHD Grant HD01994 to Haskins Laboratories. I am grateful to Quentin Summerfield for comments on an earlier version and for serving as a subject in Experiment I. Requests for reprints should be sent to Dr. Bruno Repp, Haskins Laboratories, 270 Crown St., New Haven, Conn. 06511.

(1) To demonstrate large right-ear advantages and reliable individual differences in ear dominance. (See points 3 and 4 below.) These were main objectives: Although the preceding paragraphs emphasized methodological considerations, the present studies were equally motivated on theoretical grounds.

(2) To demonstrate that similar large effects should be obtained with stimuli that are acoustically similar to the VOT stimulus. It is the difference in ear dominance for voicing contrasts in a single-response paradigm. This was an attempt to replicate the extraordinary large effects obtained by Repp (1977a) with double-feature contrasts. If it is the voicing difference that results in a resulting partial fusion of the stimuli that is responsible for the large effects, similar large effects should be obtained with stimuli that are acoustically different.

(3) To investigate the rule by which combining VOTs are fusioned to detail and to the range over which changes are possible. This also amounted to an extension of Repp (1977a), with more stimuli. The systemically changed by varying the VOTs of the components in be systematically stimulus dominance relationships demonstrated that intonational relations.

ased"; i.e., the perfectly used vowel portion is preceded by an unused portion of very brief duration. This unused portion results from aspiration noise in one ear (the initial portion of the voiceless stimulus) being compensated by a different noise and/or a periodic waveform in the other ear (the initial portion of the voiced stimulus) as long as the voice onset time (VOT) of the voiceless stimulus, perhaps 50 msec. The reciprocal result is a single stimulus accompanied by some brief noise one or the other ear. Thus, the single-response paradigm is appropriate for such partially used stimuli. Repp (1977a) showed that stimulus dominance can be varied by changing the acoustic structure of the stimuli (particularly their VOTs). The only problem with double-freeze contrasts is that they lead to a large number of "blended" responses, to four responses each stimulus combination (cf. Cutt, 1976). Blended responses are a kind of higher-level fusion and convey direct information about ear dominance. When calculating ear dominance indices, either a large amount of the data must be discarded or separate ear dominance indices must be calculated for the voicing and voice indices, either a large amount of the data must be discarded or voiceless dimensions, which raises methodological problems, since these indices are often not equal (Repp, 1977a).

The set of stimuli used most frequently in recent dichotomic studies is composed of the consonant-vowel syllables /ba/, /da/, /ga/, /ta/, and /ka/. The fifteen possible combinations of these syllables fall into three sets: place contrasts (*b/a-/d/a*, *b/a-/g/a*, *p/a-/t/a*, *p/a-/k/a*, and /t/a-/k/a/), voicing contrasts (*b/a-/p/a*, *d/a-/t/a*, and *g/a-/k/a*), and double-feature contrasts (*b/a-/t/a*, *b/a-/k/a*, *d/a-/p/a*, */g/a-/k/a*, and /p/a-, and /g/a-/t/a/). Dichotomic (voiced) place contrasts have been investigated in detail by Reppl (1976b). When precisely aligned and minimally distinictive synthetic syllables are used, these stimuli use perfectly and virtually indistinguishable from bimaural syllables. This makes them ideal for the single-response paradigm. However, they yield only small ear advantages; perfect fusion seems to prevent strong lateral asymmetries. Some of the place contrasts yield a third response category ("psychoacoustic fusion", cf. Cutting, 1976). Although stimulus dominance relations have shown that some stimulus combinations have extreme interactions that are difficult to remove. Thus, place contrasts (voiced place bases that are least, voiceless place contrasts have not been investigated in detail) are problematic with regard to three of the requirements named above.

Double-feature contrasts were investigated in detail by Reppl (1977a). Precisely aligned synthetic syllables contrasting only in the relevant acoustic parameters (voice onset time, formant transitions) are "parallelly" aligned in detail; voiceless contrasts place contrasts have not been investigated in detail; place contrasts (voiced place bases that are most difficult to remove. Thus, place contrasts (voiced place bases that are least, voiceless place contrasts have not been investigated in detail) are problematic with regard to three of the requirements named above.

error. In order to satisfy this (testable) assumption, the set of stimuli used in a test should be as homogeneous as possible in terms of their phonetic and auditory properties. Third, each stimulus combination should receive only two different categories of responses reflecting perception dominance of one or the other component stimulus. Stimulus combinations of one or the other component stimulus should generate more than two distinct responses at higher-level fusion and generate more than two distinct responses at the level of individual subjects to bring very different biases to a task, and a wide variation of intrinsic (i.e., average or expected) stimulus dominance relationships makes it less likely that some individuals show such strong biases in most stimulus pairs than others. Fourth, stimulus dominance relationships permits the actual derivation of an ROC function, whose shape determines the index of sensitivity to be used. Fifth, for the whole effort to be worthwhile, the stimuli used must generate reliable asymmetries and individual differences in ear dominance. We assume that ear dominance varies in degree between individuals and can be measured on (at least) an ordinal scale (cf. Shankweiler & Studdert-Kennedy, 1975).

perceptually integrated. Repp (1977a) obtained a curious interaction: When the voiceless stimulus in one ear had a VOT of +40 msec, a change in the VOT of the voiced stimulus in the other ear from 0 to +15 msec reduced the percentage of voiced responses, as expected; however, when the voiceless stimulus had a VOT of +55 msec, the same change in the voiced stimulus had a slight effect in the opposite direction. The present studies attempted to clarify this interaction by using more steps on the VOT dimension and by factorially combining different VOTs of voiced and voiceless stimuli. Assuming that the interaction would no longer be obtained or could otherwise be accounted for, the question may be asked: According to what rule are competing VOTs integrated into a single percept? Can the process be described by an additive or by a multiplicative model? The theory of functional measurement provides an appropriate framework for this purpose (Anderson, 1974; Massaro & Cohen, 1976).

(4) To investigate the shape of the ROC (isolaterality) function connecting points of equal ear dominance at different levels of stimulus dominance. Since variations in stimulus dominance were to be produced by varying VOT only, and since the place feature was held constant in each test, the stimuli were maximally homogeneous, and constancy of ear dominance with changes in VOT could be safely assumed. Nevertheless, this assumption could be tested by examining the scatter of the data points representing individual stimulus pairs. If these points do not lie on any single smooth function, lack of homogeneity would be indicated. If they do, the shape of the function would be of great theoretical and practical interest. Repp (1977b) proposed an index of ear dominance (called  $e'$ ) based on the assumption that the data points (when plotted as "hits" against "false alarms," as described below) follow a linear function that passes through the origin of the unit square, i.e., a linear approximation to the standard ROC function of signal detection theory. The  $e'$  index assumes values between +1 (maximal right-ear dominance) and -1 (maximal left-ear dominance). It was first applied by Repp (1977a), and the data of that experiment supported the assumptions, although there was considerable variability in the data (see Repp, 1977b, Fig. 4). The present studies provided an opportunity for further testing the assumptions underlying the  $e'$  index.

## EXPERIMENT I

### Method

#### *Subjects*

The subjects were eight Yale undergraduates, paid volunteers, some of whom had participated in earlier experiments using synthetic speech and dichotic listening. In addition,

the author and a colleague, both highly experienced listeners, participated. Two of the less experienced subjects considered themselves to be left-handed.

### *Stimuli*

The stimuli were generated on the Haskins Laboratories parallel resonance synthesizer and were similar to those used in Repp (1977a). All syllables were 300 msec long, had no initial bursts, and had a constant fundamental frequency (90 Hz). Different VOTs were generated by setting the amplitude of the first formant to zero and exciting the higher formants with a random noise source for the time specified. There were eight different VOTs: four appropriate for voiced consonants (0, +5, +10, and +15 msec) and four appropriate for voiceless consonants (+40, +45, +50, and +55 msec).

There were two parallel series of stimuli, one containing only labials (/ba/-/pa/), the other only velars (/ga/-/ka/). The stimuli were digitized (with a random sampling error of 0.125 msec) and recorded on tape using the Haskins Laboratories PCM system. Each dichotic series was preceded by 80 binaural syllables, a randomized sequence of the eight stimuli repeated 10 times. Each dichotic sequence contained 320 stimulus pairs: the 16 VOT combinations (four voiced stimuli paired with four voiceless stimuli) in the two possible channel assignments, in 10 successive individually randomized blocks of 32. The dichotic stimuli were onset-aligned, with a maximal error of 0.125 msec. The interstimulus interval was 3 sec.

### *Procedure*

The tapes were played back on an Ampex AG-500 tape recorder, and the subjects listened over Telephonics TDH-39 earphones. The intensities of the two channels were carefully equalized at approximately 75 dB SPL (peak deflections on a voltmeter). The channels were reversed electronically between the labial and velar stimulus series whose order was counterbalanced across subjects. The subjects were not informed about the nature of the dichotic stimuli. Their task was to rate each stimulus heard on a six-point scale. Ratings 1 to 3 signified /ba/ or /ga/ (1, a clear instance; 3, ambiguous but more like /ba/ or /ga/); Ratings 4 to 6 signified /pa/ or /ka/ (4, ambiguous but more like /pa/ or /ka/; 6, a clear instance). It became obvious in the analysis that the ratings did not provide any significant information beyond that obtained from simply collapsing the ratings into two categories, voiced (1-3) and voiceless (4-6). Therefore, the results are reported here in terms of percentages of voiced responses.

## Results

### *Stimulus Dominance*

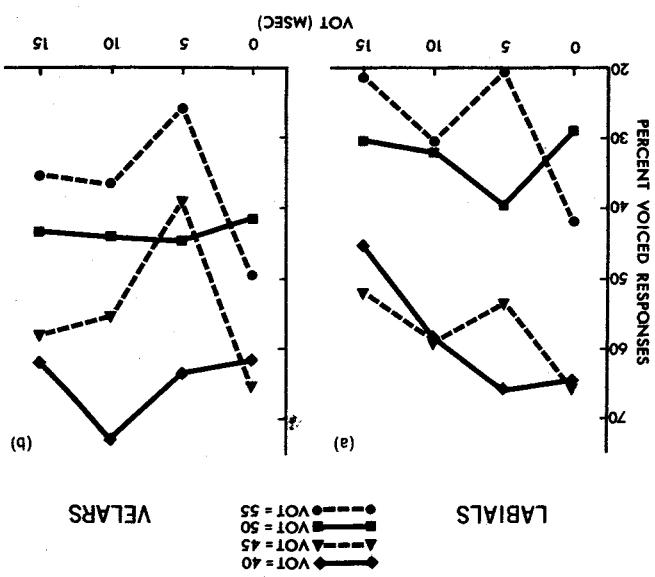
The stimuli were all reliably categorized in isolation (binaural presentation). The percentages of voiced responses are shown in Table 1. The few errors that occurred reflected the different locations of the category boundaries on the labial and velar continua. The average boundary on a labial continuum such as the present one is typically at a VOT of +25 msec, while that on a similar velar continuum occurs around a VOT of +30 msec (Miller, 1977a). Table 1 shows that the /ba/ closest to the boundary (VOT = +15 msec) received some /pa/ responses, while the corresponding /ga/ was consistently identified. On the other hand, the

At first, these results seemed extremely puzzling. The significant effects obtained showed that the interqualities were not just random variation, and it was also obvious that alternative pairs of functions in Fig. 1 showed a certain parallelism, although it was not clear why. There must have been some uncontrollable factor in the experiment that influenced stimuli dominancy. Although Reppl (1977a) also obtained an interaction between the two VOT effects, the precise pattern of this interaction was not replicated, which added to the confusion.

Eventually, the solution was found in the relative phase of the periodic portions of the dichotomic stimuli. The present stimuli had a fundamental frequency of 90 Hz, so that one period of the waveform lasted 1.1 msec. The periodicity began at the VOT specified and continued with a new pulse occurring every 1.1 msec. Since the VOTs were specified in 5-msec steps, the periodic waveforms of some stimuli combinations were nearly in phase, while others were completely out of phase. For example, the fourth and fifth pitch pulses of a stimulus with a VOT of 0 occurred at 44.4 and 55.5 msec, respectively. Thus, this stimulus was nearly in phase with stimuli whose VOTs were +40 or +50 msec. Table 2 shows the phase relationships for the different stimulus combinations.

It is clear from Table 2 that stimuli whose VOTs differed by a multiple of 10 were out of phase (the maximal pulse asynchrony being 5.5 msec), while the remaining stimuli combinations were more or less in phase.

FIG. I. Percentages of vocal



This expectation was borne out: overall, labials received 44.4% voiced responses and velars 50.0%. However, this effect did not reach significance because two subjects showed a difference in the opposite direction and because none at all. The overall percentages of voiced responses ranged from 25.1 to 63.0%. The effects of the variations in VOT are shown in Fig. 1, separately for labials and velars. The four functions correspond to the four different VOTs of voiceless stimuli; the VOTs of the voiced stimuli are on the abscissa. Thus, the vertical separation between the functions represents the effect of varying the VOT of the voiceless components, while deviations from horizontality represent the effect of varying the VOT of the voiced component. It was expected that the functions would be well separated (with the longest VOT at the bottom), monotonically decreasing from left to right, and parallel (if the two VOT effects are linearly independent and an additive model applies).

The functions obtained were remarkably irregular. Clearly, the VOT of the voiceless stimulus had a pronounced effect ( $F(3,27) = 48.6, p < .01$ ), but was nevertheless significant ( $F(3,27) = 6.5, p < .01$ ). The four functions were far from parallel, which was reflected in a highly significant interaction of the two VOT effects ( $F(9,18) = 6.7, p < .01$ ). In addition, each VOT effect interacted with the place (labial vs. velar) factor [ $F(3,27) = 3.0, p < .05$ ;  $F(3,27) = 7.5, p < .01$ ].

#### PERCENTAGES OF VOICED RESPONSES TO THE STIMULI PRESENTED BINARILY

The data of Fig. 1 are replotted in Fig. 2, separately for in-phase and out-of-phase stimulus pairs. To save space, labials and velars have been combined in this figure.

The transformation of the irregular pattern of Fig. 1 into the orderly pattern of Fig. 2 is quite remarkable. In particular, it turned out that once the data were partitioned according to phase the four functions were nearly parallel within each set of data. In-phase and out-of-phase pairs were analyzed separately in four-way analyses of variance. Three of the factors were place (labial vs velar), VOT of voiced stimulus (0 and +5 msec vs +10 and +15 msec), and VOT of voiceless stimulus (+40 and +45 msec vs +50 and +55 msec). The fourth factor ("VOT shift") represented the effect of simultaneous 5-msec changes in the VOTs of both stimuli, the difference between the solid and the dashed functions in Fig. 2.

For in-phase pairs, there was a significant decrease in the percentage of voiced responses as the VOT of the voiced stimulus increased [ $F(1,9) = 11.9, p < .01$ ] and an even larger decrease as the VOT of the voiceless stimulus increased [ $F(1,9) = 36.3, p \ll .01$ ]. The interaction between the two factors was far from significance, which confirms the parallelism of the functions in Fig. 2a. The main effect of VOT shift was not significant either. Note that this factor represents here simultaneous 5-msec changes in the VOTs of the two stimuli in *opposite* directions. Thus, the two VOT effects canceled each other in this case, despite the fact that, otherwise, VOT changes in the voiceless stimulus had a larger effect than VOT changes in the voiced stimulus. The only other significant effect was a triple interaction between place, VOT shift, and VOT of the voiced stimulus [ $F(1,9) = 8.8, p < .05$ ]. This was due to the fact that, in velars, a change in VOT from 0 to +10 msec resulted in a decrease in voiced responses, but a change from +5 to +15 msec did not; in labials, the pattern was reversed, if anything (cf. Fig. 1).

TABLE 2

PHASE RELATIONSHIPS BETWEEN THE PERIODIC STIMULUS PORTIONS  
IN ALL DICHOTIC COMBINATIONS<sup>a</sup>

VOT (msec)	+40	+45	+50	+55
0	-4.4 <sup>b</sup>	0.6	-5.5	-0.5
+5	1.7	-4.4	0.6	-5.5
+10	-3.3	1.7	-4.4	0.6
+15	2.8	-3.3	1.7	-4.4

<sup>a</sup> Temporal asynchronies between pitch pulses (phase relationships) are given in milliseconds.

<sup>b</sup> Negative values indicate that the first pulse of the voiceless stimulus preceded the temporally closest pulse of the voiced stimulus.

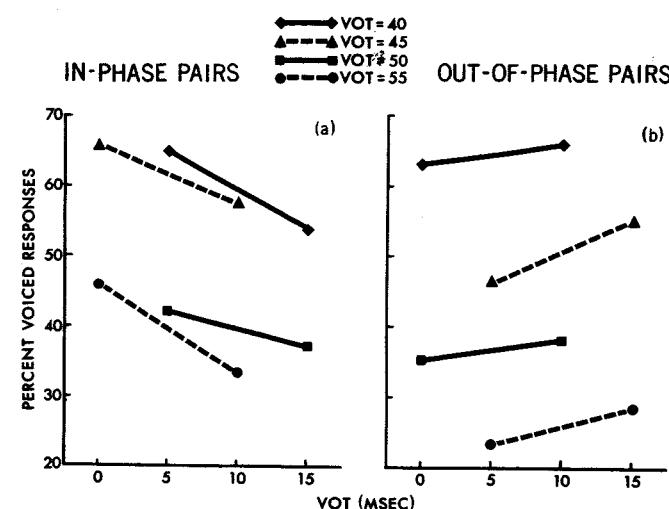


FIG. 2. Percentages of voiced responses to (a) in-phase and (b) out-of-phase stimulus pairs.

For out-of-phase pairs, there was a significant effect of the VOT of the voiced stimulus [ $F(1,9) = 17.9, p < .01$ ], but, surprisingly, it went in the opposite direction: The percentage of voiced responses *increased* with the VOT of the voiced stimulus! The effect of the VOT of the voice-

TABLE 3  
EAR DOMINANCE COEFFICIENTS FOR INDIVIDUAL SUBJECTS IN THE TWO TESTS

Subject	$e'$	
	Labials	Velars
1 <sup>a</sup>	+0.89	+0.87
2	+0.44	+0.08 <sup>b</sup>
3	+0.54	+0.96
4 <sup>c</sup>	-0.59	+0.01 <sup>b</sup>
5 <sup>a</sup>	+0.70	+0.71
6	+0.74	+0.75
7	+0.15 <sup>e</sup>	+0.62
8	+0.54	+0.59
BHR <sup>d</sup>	+0.95	+0.64
AQS <sup>e</sup>	+0.65	+0.86

<sup>a</sup> Left-handers.

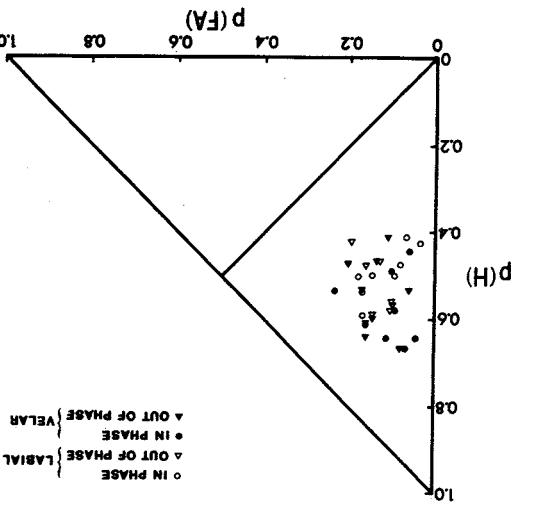
<sup>b</sup> Not significant. All other coefficients are significant at  $p < .05$  or better (see Repp, 1977b, for procedure).

<sup>c</sup> This subject's coefficients may have incorrect signs (see text).

<sup>d</sup> The author.

<sup>e</sup> A colleague.

HIC, 3. Bar dominance for 32 individual stimulus pairs in Exp 1, plotted as "hits" against "false alarms."



The ROC function must be symmetric around the negative diagonal because of the complementarity of the two response categories, each of which may be divided into hits and false alarms. Therefore, it was sufficient to consider only the less-frequent response to each stimulus pair and thus only the area below the negative diagonal of the unit square pair (Repp, 1977b). For example, if voiceless responses were less frequent than voiceless responses for a given stimulus pair, then voiceless responses given when the voiceless stimulus was in the right ear constituted hits.

The present experiment offered an opportunity to test the assumptions underlying the  $\epsilon$ , index of ear dominance as well as the assumption of test homogeneity. The homogeneity assumption says that the data points for individual stimuli combinations, when plotted as „hits”, against „false alarms”, as described below, should lie on a single, monotonic function, except for random variability. The  $\epsilon$ , index is based on the assumption that this function, the isolaterality contour or ROC function, is a linear (or slightly curvilinear) function that passes through the origin of the unit square. The unit square is the plot of the proportions of hits against the proportions of false alarms, familiar from signal detection theory.

The present tests was distinctly lower than their reliabilities. All correlations were probably somewhat overestimated due to the small number of subjects and the large between-subject variance. A more thorough evaluation of the reliability of the present tests will require a larger sample.

#### DICHOITIC VOICING CONTRASTS

The expected large right-ear advantage was obtained. The ear name coefficients for the individual subjects are shown in Table 3 separately for the two tests. It can be seen that all subjects but one showed a large right-ear dominance in at least one test. One of the two left-hand had also participated in other experiments (Repp, 1977a) and had shown large right-ear advantage there; the same is true for the author. The large left-ear advantage (subject 4) may represent a mistake in recording the channel-to-channel assignments for this subject. (He was later retested with double-feature contrasts (+0.55; Repp, 1977a). It is noteworthy, however, that 6 of the 10 subjects showed substantial differences in ear dominance between the two tests. These differences were not related to order of presentation. The correlation between two tests was +0.70 (but only +0.29 if the sign of the only left-ear advantage was reversed). This correlation does not reflect low test-reliability (Stepped-up test-reliabilities obtained by correlating ear dominance coefficients for the first and second halves of each test and subsequently applying the Spearman-Brown formula; see Lord & Novick, 1971) or the labial and velar tests were +0.99 and +0.93, respectively (or +0.95 and +0.92, respectively, if the signs of the ear dominance for the labial and velar tests were reversed). Thus, the correlation between

Ear Dominic

less stimuli was in the expected direction and highly significant ( $F = 63.1, p < .01$ ), and so was the effect of VOT shift ( $F(1,9) = 35.9, p < .001$ ). Here, VOT shift represented simultaneous 5 msec changes in the VOT of the two stimuli in the same direction. The inverted effect of the VOT of the two stimuli was apparently not strong enough to cancel the two VOT effects of the voiceless stimuli. Again, the interaction between VOT and the two VOT effects was far from significance, confirming the parallel nature of the two VOT effects in Fig. 2b. There was a highly significant interaction between place and VOT ( $F(1,9) = 22.2, p < .01$ ) and a marginally significant interaction between place and VOT ( $F(1,9) = 5.3, p < .05$ ). The first interaction resulted from the VOT effect at all. The other interaction was negligible.

Thus, the primary effect of phase was on the effect of the VOT. The voiced responses than out-of-phase pairs (cf. Fig. 2).

The voiced stimulus, in addition, in-phase pairs generally received more effect than out-of-phase pairs (cf. Fig. 1); the factor had no system to the verbal stimuli (cf. Fig. 1); in labial stimuli, the factor had no effect on the inverted effect of the VOT of the voiced stimulus was entirely ( $F(1,9) = 5.3, p < .05$ ). The first interaction between place and VOT ( $F(1,9) = 22.2, p < .01$ ) and a marginally significant interaction between place and VOT ( $F(1,9) = 5.3, p < .05$ ) resulted from the VOT effect at all. The other interaction was negligible.

and voiceless responses given when the voiceless stimulus was in the left ear constituted false alarms. These proportions were averaged over all subjects (excluding subject 4) for each individual stimulus combination, so that 32 data points were obtained. These are plotted in Fig. 3.

The results were disappointing. The 32 data points clustered in a roughly circular area in the left-hand quadrant of the unit square. The stimuli were homogeneous in so far as all showed sizeable average right-ear advantages. However, variability was so large that it was impossible to determine a single function that fitted the point swarm well. The variation was not systematically related to either the place distinction or the phase factor. All that can be concluded is that there was large, probably random, variation between stimulus pairs. For a critical test of the assumptions underlying the  $e'$  index, either more observations per stimulus pair or more extreme stimulus dominance relationships are needed.

### Discussion

The present study achieved several of its goals. It showed that stimulus dominance relationships in voicing contrasts can be systematically changed by varying the VOTs of the component stimuli, particularly the VOT of the voiceless stimulus in a pair. After taking the phase factor into account, it became clear that the perceptual integration of the VOTs of the two stimuli was approximately linear and additive. Most subjects showed extremely large right-ear advantages, which replicated the findings of Repp (1977a). Only the question of item homogeneity and the shape of the isolaterality contour remained undecided, but at least the results did not directly contradict the assumptions underlying the  $e'$  index of ear dominance.

A puzzle was created by the unexpected effect of the relative phase of the stimuli. Why did phase have any effect at all? Why was the effect of the VOT of the voiced stimulus reversed when the stimuli were out of phase? And why did out-of-phase stimuli receive fewer voiced responses? Although phase may be expected to have some effect on fusion, there was no indication that in-phase and out-of-phase stimulus pairs were phenomenologically different. To the author, the test sequences seemed perceptually quite homogeneous, and no stimulus pairs sounded less fused than others.

Note that relative phase applied only to the simultaneous periodic portions of the stimuli, i.e., the vocalic portions *after* voicing onset in the voiceless stimulus. Therefore, it was especially surprising that phase changed the effect of the VOT of the voiced stimulus, since this voicing onset occurred at a time when phase could not yet have played a role. One way of describing the results would be that the voicing feature of in-phase stimuli was determined by a weighted average of the VOTs

of the two stimuli, but, in out-of-phase stimuli, it was determined by the weighted *difference* of the two VOTs. If the difference between the two competing VOTs played a role, an inverted effect of the VOT of the voiced stimulus would be expected, as well as a relative decrease in voiced responses. Perhaps the decision mechanism responsible for the voicing feature is sensitive to the intervals between *any* abrupt changes in energy at and shortly after stimulus onset. Normally, there are only two such energy increments: stimulus onset and voicing onset. In the present partially fused stimuli, however, there were three: stimulus onset, voicing onset in the voiced stimulus, and voicing onset in the voiceless stimulus. Thus, there were three temporal intervals, and the probability of hearing a voiced consonant may have been a weighted function of all three. That intervals other than VOT can affect voicing judgments was demonstrated by Repp (1976a). He found that the interval between stimulus onset in one ear and the onset of an isolated vowel in the other ear had a significant influence on the probability of voiced responses, in addition to VOT. However, there was no indication in these data that the interval between voicing onset in one ear and vowel onset in the other ear played a role, although the relative phase of the dichotic stimuli varied, as in the present experiment. Thus, the phase effects obtained here remain unexplained.

Intriguing as the phase effect was, it was basically an ancillary finding and was not essential to the theoretical and methodological purpose of the present experiment. Therefore, it was decided not to investigate the effect further, for the time being, but instead to attempt to replicate Expt I with stimuli that were definitely in phase. (Even the in-phase stimuli of Expt I were slightly out of phase.) This was achieved in Expt II by choosing the VOTs of the voiceless stimuli so that they coincided with pitch pulses in the voiced stimuli.

## EXPERIMENT II

### Method

#### Subjects

The subjects were seven new paid volunteers and the author. Again, the subjects had had varying degrees of exposure to synthetic speech. One subject was left-handed (self-report). The data of one additional subject were rejected because they were too variable.

#### Stimuli

The stimuli for this experiment were generated on the OVE IIIC synthesizer at Haskins Laboratories, a serial resonance synthesizer that permitted finer control of certain stimulus parameters and tends to produce somewhat more "natural" speech. This time, a continuum

## DICHOTIC VOICING CONTRASTS

PERCENTAGES OF VOICED RESPONSES TO BINAURAL STIMULI AND DICHOTIC WITHIN-CATEGORY COMBINATIONS						
VOT (msec)	0	+8	+16	+24	+32	+40
0	100.0	100.0	100.0	100.0	95.0	51.3
+8	100.0	100.0	100.0	100.0	98.8	99.4
+16	100.0	100.0	100.0	100.0	100.0	100.0
+24	99.4	98.8	95.0	51.3	0.0	0.0
+32	0.0	0.0	0.0	0.0	0.0	0.0
+40	1.3	2.5	0.0	0.0	0.0	0.0
+48	4.0	1.9	1.9	1.9	1.9	1.9
+56	5.6	0.0	0.0	0.0	0.0	0.0

subject whose data were somewhat noisy. The most interesting result in Table 4 is that there were hardly any voiceless responses to combinations of stimuli within the voiced category. The results for the between-category combinations are shown in Fig. 4. The data were very orderly and confirmed the predictions. The effect of the VOT of the voiced stimulus was highly significant [ $F(3,21) = 47.9$ ,  $p \leq .01$ ], and so was the effect of the VOT of the predicted interaction [ $F(3,21) = 38.5$ ,  $p \leq .01$ ]. In addition, there was a significant interaction between the two factors [ $F(9,63) = 11.4$ ,  $p \leq .01$ ]. Likewise, the voiceless stimulus closest to the boundary ( $VOT = +24$  msec), strongly dominated not only by voiced but also by voiceless stimuli. The interaction was due to the  $VOT = +24$  msec stimulus, which was stronger than the two factors [ $F(9,63) = 11.4$ ,  $p \leq .01$ ].

The results for the between-category combinations are shown in Fig. 4. The data were very orderly and confirmed the predictions. The effect of the VOT of the voiced stimulus was highly significant [ $F(3,21) = 47.9$ ,  $p \leq .01$ ], and so was the effect of the VOT of the predicted interaction [ $F(3,21) = 38.5$ ,  $p \leq .01$ ]. In addition, there was a significant interaction between the two factors [ $F(9,63) = 11.4$ ,  $p \leq .01$ ]. The more strongly dominated not only by voiced but also by voiceless stimuli, the stronger was the interaction between the two factors [ $F(9,63) = 11.4$ ,  $p \leq .01$ ].

In Table 4, all other coefficients are significant at  $p < .05$  or better.

VOT (msec)	0	+8	+16	+24	+32	+40	+48	+56
0	100.0	100.0	100.0	100.0	95.0	51.3	0.0	0.0
+8	100.0	100.0	100.0	100.0	98.8	99.4	0.0	0.0
+16	100.0	100.0	100.0	100.0	100.0	100.0	0.0	0.0
+24	99.4	98.8	95.0	51.3	0.0	0.0	0.0	0.0
+32	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
+40	1.3	2.5	0.0	0.0	0.0	0.0	0.0	0.0
+48	4.0	1.9	1.9	1.9	1.9	1.9	1.9	1.9
+56	5.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0

TABLE 4

PERCENTAGES OF VOICED RESPONSES TO BINAURAL STIMULI AND DICHOTIC WITHIN-CATEGORY COMBINATIONS

Most of the errors in voiceless stimuli pairs stemmed from a single trial. The percentages of voiceless responses to the eight binaural stimuli and to the within-category dichotic pairs are shown in Table 4. It can be seen that all stimuli but one were identified with high consistency. The stimulus with  $VOT = +24$  msec fell just about at the average category boundary. The percentage of voiceless responses to the eight binaural stimuli and to the within-category dichotic pairs are shown in Table 4. It can be seen that the within-category dichotic pairs are identified with high consistency. The stimulus with  $VOT = +24$  msec fell just about at the average category boundary.

## Stimulus Dominance

## Results

strongly that their responses were not contingent on where they heard the aspiration noise. To the fused stimuli in the middle of the head, and the subjects suggested in which ear this stimulus occurred on a given trial. Nevertheless, it was easy to attend what more noticeable than in Exp I; if required, it would not have been difficult to tell more noise than in Exp I, about 85 dB SPL (peak deflections on a voltmeter). The more acoustic segregation of the syllables they heard. All subjects were told to ignore any noises simply told to identify the dichotic nature of the stimuli. In the present stimuli, the experience than in Exp I, about 85 dB SPL (peak deflections on a voltmeter). The more channels were reversed electronically before the repetition. Stimulus intensity was twice. Channels were subject to the binaural stimuli, each subject listened to the dichotic series twice. After listening to the binaural stimuli, each subject responded by writing down Dr T. After listening to the binaural stimuli, each subject listened to the dichotic series twice. Channels were reversed electronically before the repetition. Stimulus intensity was twice. Channels were reversed electronically before the repetition. Stimulus intensity was twice. The rating scale was no longer used; the subjects simply responded by writing down

## Procedure

The rating scale provided by writing down extreme precision. The intersimulus interval was 3 sec. by providing unambiguous "analogous" stimuli. The dichotic stimuli were onset-aligned with contrast to Exp I, within-category combinations of the eight stimuli, in both channel assignments. In all possible dichotic combinations of 56 dichotic pairs followed. Each block contained 10 times. A sequence of five blocks of 80 binaural syllables—the 8 stimuli replicated 10 times. The stimulus tape contained first a 10-KHz sampling rate (time-locked to stimulus onset). The stimulus tape digitized at a 10-KHz sampling rate (time-locked to vocalic portion. The effective amplitude of the aspirated portion was still below that of the vocalic portion. generalization. However, since the two parameters were not on the same scale, greater variance was selected to be 5 dB higher than the subsequent amplitude setting for the noise generated. In the syllables specific conditions, the amplitude setting for the noise generated. In the syllables specific conditions, the amplitude setting for the noise generated. (which can be controlled in the OVE synthesizer); if the amplitude is too low, dichotic influence dichotic stimulus dominance suggests that an additional factor exactly at the VOT specified. The pulse generator of the first genuine pitch pulse occurred synthesizer was synchronized to stimulus onset, so that the first genuine pitch pulse occurred the first genuine voice pulse had the intended amplitude. The pulse generator of the kept at minimum amplitude during the aspirated portion of the signal; this ensured that kept to its maximum bandwidth. The pulse generator was turned on at stimulus onset but maintained to its maximum bandwidth. The pulse generator excitation and setting the first for-gory but fell in the region of the phone me boundary.

three were equal, the fourth  $VOT = 0$ , +8, +16, +24, +32, +40, +48, and +56 msec. Because of the wider spacing, the fourth  $VOT = +24$  msec) was no longer entirely within the voiced category. All stimuli were spaced according to frequency in 8 msec steps. As in Exp I, period of 8 msec, and the  $VOTs = 0$ , +8, +16, +24, +32, +40, +48, and +56 msec. Because of the wider spacing, the fourth  $VOT = +24$  msec) was no longer entirely within the voiced category. All stimuli were spaced according to frequency in 8 msec steps. As in Exp I, 300 msec long and had a constant fundamental frequency of 125 Hz. This resulted in a 300 msec long and had a constant fundamental frequency of 125 Hz. All stimuli were of a vowel-like stops was selected, ranging periodically from /da/ to /ta/. All stimuli were

although consistently identified in isolation, was strongly dominated by the voiced stimuli and did not completely dominate the VOT = +24 msec stimulus. Thus, both stimuli flanking the boundary were weak in dichotic competition. The data were reanalyzed omitting these two stimuli. When only combinations of "good instances" of each category were considered, the effect of the VOT of the voiced stimulus was much reduced but was nevertheless in the predicted direction and was significant [ $F(2,14) = 7.3, p < .01$ ], the effect of the VOT of the voiceless stimulus was more pronounced [ $F(2,14) = 15.5, p < .01$ ], and there was no longer any significant interaction. In other words, the functions were again parallel (cf. Fig. 4).

#### Ear Dominance

The individual  $e'$  coefficients for the eight subjects are shown in Table 5. The average right-ear advantage in this test was smaller than in Expt I but still substantial ( $e' = +0.37$ , based on average scores). Only two subjects and the author showed very large right-ear advantages. Of the remaining subjects, four showed small right-ear advantages and one a moderate left-ear advantage. The reliability of this test was estimated by the split-half method to be +0.96, although some subjects showed considerable variation. The reliability is again somewhat overestimated, due to the small subject sample, but it is nevertheless encouraging.

Figure 5 shows the average hit and false alarm proportions for the

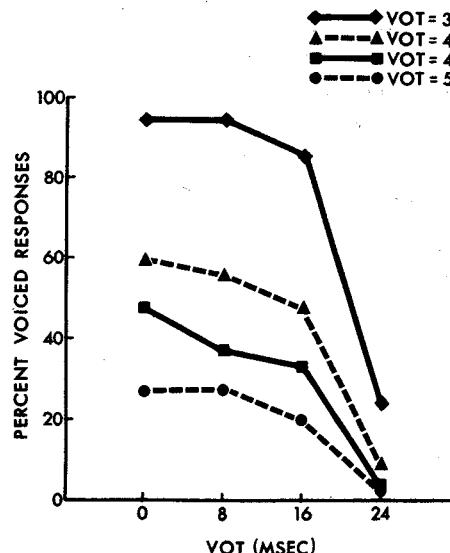


FIG. 4. Percentage of voiced responses to sixteen VOT combinations in Expt II.

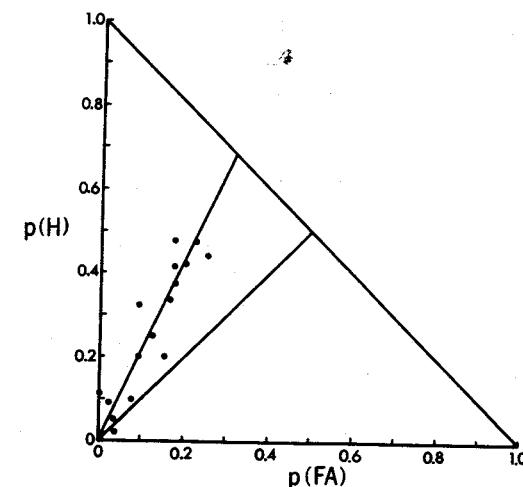


FIG. 5. Ear dominance for 16 individual stimulus pairs in Expt II, plotted as in Fig. 3, with the best-fitting linear isolaterality contour drawn in.

sixteen stimulus pairs. It can be seen that, in this test, much more variation in the average "bias" was obtained than in Expt I, a consequence of including stimuli close to the category boundary. The points are clearly best fitted by a linear (or perhaps curvilinear) function through the origin and not, for example, by a linear function parallel to the positive diagonal. This latter function would be the isolaterality contour corresponding to the simple difference score,  $p(H) - p(FA)$ , as an index of ear dominance (cf. Repp, 1977b). The present data strongly argue against this simple difference index and support the assumptions underlying the  $e'$  index.

#### Discussion

Experiment II successfully replicated the results obtained in the in-phase condition of Expt I. The competing VOTs of voicing contrasts appear to be perceptually integrated according to an additive rule, as long as neither VOT is too close to the category boundary. When the VOT of one of the competing stimuli approaches the category boundary, this stimulus loses competitive strength and is dominated by the opponent stimulus. Changes in the VOT of the voiceless stimulus have a more pronounced effect than changes of equal magnitude in the VOT of the voiced stimulus.

The average right-ear advantage in the present experiment was not as large as in Expt. I and in earlier results (Repp, 1977a). However, the exceptionally large average effects obtained earlier were probably fortuitous and due to the small subject samples. These earlier tests probably just happened to include a number of subjects from the upper end of

not so strongly fused as to suppress lateralization effects (cf. Repp, 1976b). Obviously, the small difference between the two competing stimuli in their first 40 to 60 msec is sufficient to produce strong ear asymmetries. It is intriguing to speculate that there is a direct relationship between dichotic fusion and the suppression of ipsilateral auditory transmission, one of the factors responsible for the ear advantage according to Kimura's (1961) original theory. The auditory discrepancy at the onset of dichotic voicing contrasts (periodic vs noise excitation) may lead to very effective ipsilateral suppression in this connection that results from the early lateralization effect (Repp, 1977a). Research is now in progress to determine whether dichotic lateralization tests and different methodologies assess the same factor of laterality or if perhaps multiple factors are involved.

Clearly, there is still much to be learned about measuring ear advantage. Ingle (1977a) has found different tests and different methods to be correlated with each other (Repp, 1977a). Research is now in progress to determine whether dichotic lateralization tests and the voice leading and place measures found earlier (Repp, 1977a) are correlated with each other. The question is also raised for which reason this finding is reminiscent of a similarity of also reason for concern. This finding is reminiscent of a similarity of relative low correlation between the lateral and verbal tests in Experiment I. The same phenomenon that traditional two-response tests measure, the same stages, it is necessary to ask whether the present tests measure advantages, what is their validity? Because of the unusual magnitude of the ear advantage, it is difficult to control.

The principal question remains: What do the present tests measure, a permanent noise factor that is difficult to control, two separate events are heard and selective-attention strategies constitute by proper instructions, unlike the situation with unfused stimuli in which a separate stimulus. It seems that such differences can be avoided of a separate stimulus was in the same ear, since there is no clear percept of voiceless stimuli from this observation that the listener would have to infer from the aspiration noise occurred, the listener would tell in which ear the aspiration noise occurred though it may be possible to another methodological plus. Although it is also likely that dichotic voicing contrasts are not sensitive to selective-attention effects, which constitutes another methodological plus. Although it may be possible to tell in which ear the aspiration noise occurred, the listener would have to infer from this observation that the voiceless stimuli was in the same ear, since there is no clear percept of a separate stimulus with unfused stimuli in which, of course, has no particular phase relationship.

## REFERENCES

- Amerson, N. H., 1974. Allegorical models in perception. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception*. New York: Academic Press. Vol. 2, pp. 215-298.

Amerson, N. H., 1973. Dicoholic speech perception: An interpretation of height-eer advantage. *Journal of the Acoustical Society of America*, 53, 699-709.

Amrey, E., & Speaks, C., 1976. Effect of acoustic variation in stimuli on response patterns in dichotic listening. *Journal of the Acoustical Society of America*, 59 (Suppl. No. 1), 96 (Abstr.).

The finding that a change in the VOT of the voiced stimulus had a smaller effect than a change in the VOT of the voiceless stimulus may have been due to the fact that the stimuli increased in amplitude over the first 30 msec or so, which is the region of short VOT values. If this explanation is correct, the effect might constitute additional evidence for interaction at an auditory level. According to informal observations, the intensity of the aspiration noise is another auditory factor affecting the interaction at the VOT level. Strong dominance of voiceless stimuli dominates. Strong dominance of voiceless stimuli over voiced stimuli has sometimes been reported in the literature (e.g., Berlin, Lowe-Bell, Cullen, Thompson, & Loovis, 1973). Although these studies used natural speech stimuli, the perceptual asymmetry can almost certainly be explained in terms of the auditory properties of the stimuli (such as long VOTs and heavy aspiration of voiceless stops).

## GENERAL DISCUSSION

The distribution of large variation in early dominance is desirable for methodological purposes: The larger the variation, the more reliable will the measurements be (unless within-subject variability increases in proportion to between-subject variation).

- Cutting, J. E. 1976. Auditory and linguistic processes in speech perception: Inferences from six fusions in dichotic listening. *Psychological Review*, **83**, 114-140.
- Green, D. M., and Swets, J. A. 1966. *Signal detection and psychophysics*. New York: Wiley.
- Kimura, D. 1961. Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology*, **15**, 166-171.
- Lord, F. M., & Novick, M. R. 1968. *Statistical Theories of Mental Test Scores*. Reading, Mass.: Addison-Wesley.
- Massaro, D. W., & Cohen, M. M. 1976. The contribution of fundamental frequency and voice onset time to the /zi/-/si/ distinction. *Journal of the Acoustical Society of America*, **60**, 704-717.
- Miller, J. L. 1977a. Nonindependence of feature processing in initial consonants. *Journal of Speech and Hearing Research*, **20**, 519-528.
- Miller, J. L. 1977b. Properties of feature detectors for VOT: The voiceless channel of analysis. *Journal of the Acoustical Society of America*, **62**, 641-648.
- Repp, B. H. 1976a. Dichotic "masking" of voice onset time. *Journal of the Acoustical Society of America*, **59**, 183-194.
- Repp, B. H. 1976b. Identification of dichotic fusions. *Journal of the Acoustical Society of America*, **60**, 456-469.
- Repp, B. H. 1977a. Dichotic competition of speech sounds: The role of acoustic stimulus structure. *Journal of Experimental Psychology (Human Perception and Performance)*, **3**, 37-50.
- Repp, B. H. 1977b. Measuring laterality effects in dichotic listening. *Journal of the Acoustical Society of America*, **62**, 720-737.
- Shankweiler, D., & Studdert-Kennedy, M. 1975. A continuum of lateralization for speech perception? *Brain and Language*, **2**, 212-225.