Stimulus dominance in fused dichotic syllables: Trouble for the category goodness hypothesis

Bruno H. Repp

Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06510

Repp [J. Acoust. Soc. Am. 60, 456–469 (1976)] hypothesized that dichotic stimulus dominance relationships between fused speech sounds reflect the relative category goodness of the competing stimuli, i.e., their relative perceptual distances from the listener's linguistic category prototypes. In experiment I, 15 synthetic syllables from a /bæ/–/dæ/–/gæ/ continuum were dichotically fused with three selected stimuli from the same continuum and presented to listeners for identification. Stimuli in the vicinity of phonetic category boundaries were found to make weaker dichotic competitors than stimuli from well within a phonetic category, as predicted. However, some systematic deviations from the predictions and large individual differences were noted. In experiment II, synthetic syllables from a /ba/–/da/–/ga/ continuum were paired with either of the two endpoint stimuli (/ba/,/ga/) in dichotic or mixed presentation. The resulting fused hybrid stimuli were presented in identification and AXB discrimination tests. The hybrids were perceived quite categorically, and there was little difference between dichotic and mixed modes of presentation. These results did not replicate earlier data [Repp, Haskins SR-45/46, 123–139 (1976)] that had suggested that discrimination of dichotic hybrid stimuli might be based on a level of representation preceding phonetic categorization. The category goodness hypothesis, which implies such a level, was neither supported nor contradicted. In experiment III, the prediction was tested that increasing the aspiration amplitude of a /ta/ syllable, and thus its category goodness, would increase the perceptual dominance of this stimulus over a /da/ simultaneously presented to the other ear. This hypothesis was not supported by the data. Rather, the response pattern suggested that a listener's success in perceptually integrating the aspiration noise (of /ta/) with the fused vocalic portion (of /da/ and /ta/) into a single phonetic percept depended on the perceived spatial locations of these stimulus portions. Taken together, the results of the three experiments cast doubt on the category goodness hypothesis and suggest that dichotic stimulus dominance in fused speech sounds is determined by psychoacoustic factors, some as yet undefined.

PACS numbers: 43.70.Dn, 43.66.Rq

INTRODUCTION

In the large majority of dichotic studies using speech sounds, the focus of interest has been the direction and magnitude of the ear dominance effect (or ear advantage). However, there is a second factor that plays an important role in dichotic perception. This factor, called "stimulus dominance" (Repp, 1976a), is the tendency of one stimulus in a specific dichotic pair to receive more correct responses than the other stimulus, regardless of the ear in which it occurs. Ear dominance and stimulus dominance are independent factors that jointly determine the listener's responses to dichotic stimulus pairs.

Most dichotic experiments in the past have used stimuli that did not fuse and thus were heard as more or less separate events. Stimulus dominance effects have occasionally been noted (e.g., Berlin *et al.*, 1973), but they did not appear to be theoretically interesting at the time. More recent work using fused dichotic syllables has changed this state of affairs (Repp, 1976a, 1977a, 1977b, 1978a). Stimulus dominance plays an important role in assessing ear dominance with fused stimuli, comparable to the role of response bias in signal detection tasks (Repp, 1977b). In addition, stimulus dominance relationships may reveal some interesting facts about the nature of dichotic stimulus interaction. By identifying the properties that make one stimulus dominate another, important information may be obtained about the level at which perceptual competition between

fused dichotic stimuli takes place.

There are at least three different (not mutually exclusive) levels at which dichotic competition between speech sounds may be conceived to occur. One is the phonetic level (Studdert-Kennedy, Shankweiler, and Pisoni, 1972). If dichotic competition occurred solely between categorical phonetic representations (syllables, phonemes, or features), acoustic stimulus variations within a phonetic category should have no influence on the degree of dominance that a stimulus from that category exerts over stimuli from other categories. However, there is now ample evidence that this is not true: Within-category changes in the voice onset times or the formant transitions of competing syllable-initial stop consonants do significantly change stimulus dominance relationships (Miller, 1977; Repp, 1976a, 1977a, 1978a). Thus, dichotic competition is certainly not exclusively phonetic in nature, although there may conceivably be a phonetic component to the process.

Perhaps the most obvious level at which dichotic competition between fused syllables might occur is that of auditory processing. The phenomenon of dichotic fusion itself is determined by rather low-level auditory properties of the signals; thus, fusion can easily be prevented by introducing slight discrepancies in fundamental frequency or temporal alignment (Halwes, 1969; Cutting, 1976). However, even though fusion occurs at a relatively low level in the auditory system, it seems unlikely that the individual auditory properties of the fused

stimuli are lost at this early level by direct auditory interactions, such as masking or integration. Rather, they are probably retained in some superimposed or mixed form, and the dominance of one stimulus property over another comes about because it is, in some sense, more salient. If this salience reflects auditory stimulus properties, then we might expect stimulus dominance relationships to change (if they change at all) as a relatively smooth and continuous function of changes in acoustic stimulus parameters, as long as these changes do not introduce discontinuities in auditory perception. Let us call this the *auditory salience hypothesis.*

In experiments on dichotic voicing contrasts (syllable-initial stop consonants contrasting only in the voicing feature, such as /ba/–/pa/), this prediction has generally been supported: The voiceless stimulus in a pair becomes more (less) dominant as its voice onset time is increased (decreased); and similar, although smaller, effects are obtained by manipulating the voice onset time of the voiced stimulus in a pair (Miller, 1977; Repp, 1977a, 1978a). However, in a study of dichotic place contrasts (syllable-initial stop consonants contrasting only in place of articulation, such as /ba/–/da/), Repp (1976a) found an interesting irregularity in the effects of acoustic parameters on stimulus dominance, which led him to formulate the hypothesis that is the focus of the present studies.

In this earlier experiment, seven different consonant-vowel syllables were presented for identification in all possible dichotic pairings. The stimuli were distinguished only by the initial transitions of the second and third formants (F_2 and F_3) whose onset frequencies were varied to form a continuum from /bæ/ to /dæ/ to /gæ/. When presented dichotically, these stimuli were perfectly fused and sounded like single syllables presented binaurally. The results showed a clear effect of variations in F_2 and F_3 transitions on stimulus dominance, even when these variations occurred within a phonetic category. However, this effect was not monotonic with the changes in F_2 and F_3. In particular, when stimulus 1 (/bæ/) was paired with stimuli 3–7 (/dæ/, /gæ/), the percentage of B responses showed a local maximum in the pairing with stimulus 5, which happened to be ambiguous between /dæ/ and /gæ/. This finding suggested that yet another property of the stimuli may be important in dichotic competition: the "category goodness" of the stimuli. Stimulus 5 was ambiguous and hence not a good example of any category. This may have been the reason why it was a weaker dichotic competitor for /bæ/ than its neighbors on the stimulus continuum.

This interpretation suggests a determinant of dichotic stimulus dominance intermediate between the auditory and phonetic levels. In the context of recent models of phoneme recognition, this intermediate level has been termed the "multicategorical" (Repp, 1976a, 1977a) or "prototype matching" stage (Oden, 1978; Oden and Massaro, 1978). At this level, the perceptual system is assumed to determine how well a stimulus matches any of several category prototypes. When two competing dichotic stimuli enter the system, a stimulus that is

close to a prototype will tend to dominate over a stimulus that is far from any prototype. This is the *category goodness hypothesis.*[1] There are several possible conceptions of the mechanisms involved; assuming a "race for the nearest prototype" in auditory space would be an example.

Since variations in acoustic stimulus parameters change not only the perceptual distance of the stimuli from the category prototypes but possibly also the salience of the auditory cues, the predictions of the category goodness hypothesis for changes in stimulus dominance are often difficult to distinguish from those of the auditory salience hypothesis. Nor are the two hypotheses mutually exclusive. The evidence in favor of the category goodness hypothesis, as applied to dichotic listening, rests primarily on the demonstration that ambiguous stimuli are weak in dichotic competion with unambiguous stimuli in cases where no such weakness would be predicted on the basis of auditory salience. Repp's (1976a) results were merely suggestive in that regard. It was the purpose of experiment I to investigate this question in more detail.

I. EXPERIMENT I

In order to obtain more precise data than in the earlier study (which had used only seven different stimuli), a 15-member /bæ/–/dæ/–/gæ/ continuum was created by varying the onset frequency of the F_2 transition. (The F_3 transition was held constant, as explained below.) All stimuli were dichotically paired with themselves, and with stimuli 1, 8, and 15, which were representative of the three phonetic categories. Consider the results that might be obtained for stimulus 1 (/bæ/) when paired with all others and presented for identification as B, D, or G. When the percentage of B responses is plotted as a function of the number (i.e., location on the continuum) of the stimulus competing with stimulus 1, a "stimulus dominance function" is obtained that describes changes in the relative dominance of stimulus 1 as the formant transitions of the competing stimulus are changed. According to the auditory salience hypothesis, this function should be relatively smooth and continuous; its precise shape will depend on whether or not the other stimuli tend to dominate stimulus 1. (We assume, for the time being, that auditory salience is a smooth function of acoustic changes in formant transitions.) The category goodness hypothesis, on the other hand, predicts a significant local increase in B responses in the region of the /dæ/–/gæ/ category boundary, where the competing stimulus is ambiguous (and therefore expected to be weaker in dichotic competition). A similar stimulus dominance function may be obtained for stimulus 15 (/gæ/) paired with all others, with the percentage of G responses as the dependent variable. Here, the category goodness hypothesis predicts a local peak in the function at the /bæ/–/dæ/ boundary.

In order to make sure that these local peaks, if obtained, are really related to the category boundaries—and thus to the phonetic ambiguity of the competing stimulus—two different stimulus sets were used in the present experiment. They were distinguished by the

presence or absence of a rising F_3 transition which remained constant over the whole stimulus continuum. The F_3 transition is known to affect the perception of place of articulation, particularly the size of the alveolar category on continua such as used here (Harris *et al.*, 1958; Hoffman, 1958). Thus, the two different stimulus series were expected to have their category boundaries in different locations. Accordingly, the local peaks in the stimulus dominance functions predicted by the category goodness hypothesis should appear at corresponding different locations in the two stimulus series.

The category goodness hypothesis also predicts that a stimulus from within a given category should dominate stimuli ambiguous between this same category and a neighboring category. For example, a good /bæ/ (stimulus 1) should dominate stimuli ambiguous between /bæ/ and /dæ/. In the stimulus dominance function for stimulus 1, this should be reflected in a high level of B responses extending from the /bæ/ category beyond the /bæ/–/dæ/ boundary. Confirmation of these predictions in addition to those made above would constitute strong evidence in favor of the category goodness hypothesis.

A. Method

1. Subjects

Six subjects participated. One of them was the author; the others were paid volunteers (four Yale undergraduates and one high-school student). All but one had participated in earlier experiments and had been selected because of their accurate performance with synthetic speech stimuli. One subject had no experience with synthetic speech but did just as well as the others.

2. Stimuli

The stimuli were two sets of 15 synthetic syllables produced on the Haskins Laboratories parallel resonance synthesizer (frame rate = 200/s) and ranging perceptually from /bæ/ to /dæ/ to /gæ/. All syllables were 280 ms long, had a constant fundamental frequency of 114 Hz, a voice onset time of −15 ms (i.e., prevoicing), 45-ms stepwise-linear formant transitions including a rising transition in F_1, and no burst but an abrupt onset of energy following the prevoicing. (Very similar stimuli had been used by Pisoni, 1971, in his studies on categorical perception.) Within each set, the stimuli differed only in the onset frequencies of F_2, which are shown in Table I. One stimulus set had a constant F_3 transition that rose from 2180 to 2862 Hz ("rising F_3"). The other set had no F_3 transition; F_3 remained constant at 2862 Hz in all stimuli ("flat F_3"). The presence of a rising F_3 transition was expected to bias perception against D, and more toward B and G.

All stimuli were digitized at 8 kHz using the pulse code modulation system at Haskins Laboratories. Dichotic pairs were created from the digitized waveforms with a special-purpose computer program.[2] The tape contained three sequences of 228 stimulus pairs each.

Each sequence contained the following combinations within each stimulus set: all 15 stimuli paired with themselves twice (30 binaural pairs); stimulus 1 paired with all others, in both channel assignments (28 dichotic pairs); stimulus 8 paired with all others (28 pairs); and stimulus 15 paired with all others (28 pairs). (Pairs 1 + 8, 1 + 15, and 8 + 15 were unnecessarily duplicated and occurred twice as often as the other pairs.) Stimuli from different sets were never paired with each other. All in all, there were $30 + 3 \times 28 = 114$ pairs from each stimulus set, which were randomized together. The interpair interval was 2 s. The three sequences of 228 pairs were separated by longer intervals.

3. Procedure

The experiment required three 1-hour sessions, held on different days. The subjects were instructed to identify the initial consonants in each (fused) stimulus as B, D, or G. In each session, the stimulus tape was played twice; thus, at the end of the experiment, each subject had given a total of 36 responses to each stimulus pair (18 to each channel assignment of a dichotic pair), except for the three duplicated pairs, which received 72 responses each. The subjects knew that there were both binaural and dichotic stimuli in the sequence; but, as Repp (1976a) has shown, fused dichotic syllables contrasting only in their initial formant transitions are practically indistinguishable from binaural syllables, and the present stimulus sequences indeed sounded like homogeneous lists of single syllables originating in the center of the listener's head.

The tape was played back on an Ampex AG-500 tape recorder. The subjects listened over Telephonics TDH-39 earphones. Playback amplitude was adjusted and monitored on a Hewlett–Packard voltmeter, and special care was taken to equalize the amplitudes of the two channels at about 85 dB SPL (peak deflections on a voltmeter for a single stimulus). The tape recorder channels were reversed electronically halfway through each session, in order to counterbalance any possible quality differences between tape tracks.

TABLE I. F_2 onset frequencies of the stimuli in experiment I.

Stimulus No.	F_2 onset (Hz)
1	1155
2	1232
3	1312
4	1386
5	1465
6	1541
7	1620
8	1695
9	1772
10	1845
11	1920
12	1996
13	2078
14	2156
15	2234
Steady state of F_2	1620

B. Results

1. Binaural identification

The average labeling functions for the two sets of 15 stimuli when presented binaurally (randomized together with the dichotic pairs) are shown in Fig. 1. The solid functions represent the set with a rising F_3. It can be seen that the D category was fairly narrow, and D responses did not exceed 90% to any stimulus. This was expected, since alveolar stops normally require a falling F_3 transition. Clearly, however, an F_2 transition in the range appropriate for alveolars perceptually overrode the conflicting F_3 transition. In the stimulus set with no F_3 transition (dotted functions), the D category was more prominent and occupied a larger region on the stimulus continuum. As expected, both category boundaries shifted outward as the onset frequency of F_3 was raised (rising versus flat transition), thus reducing the frequency of B and G responses, which normally require a low F_3 onset. The shift was significant, $F(1,5) = 19.2$, $p < 0.01$. Still, the F_2 transition remained the dominant cue for the perceived place of articulation.[3]

2. Dichotic stimulus dominance functions

Figure 2 shows the stimulus dominance functions separately for the two stimulus sets (solid versus dotted lines). The results of the six subjects were averaged to obtain this figure. The left-hand panel shows the percentage of B responses for the combinations of stimulus 1 (/bæ/) with stimuli 1–15 ("+1 pairs"); the center panel shows the percentage of D responses for the combinations of stimulus 8 (/dæ/) with stimuli 1–15 ("+8 pairs"); and the right-hand panel shows the percentage of G responses for the combinations of stimulus 15 (/gæ/) with stimuli 1–15 ("+15 pairs"). The locations of the category boundaries for the binaural stimuli (Fig. 1) are indicated by arrows in Fig. 2. Ear of presentation was ignored in this analysis, and all data were collapsed over this factor. (Ear dominance results are described in the Appendix.)

Consider first the panel labeled $B(+1)$. The solid line describes the extent to which stimulus 1 (/bæ/) dominated the other stimuli on the rising-F_3 continuum, as
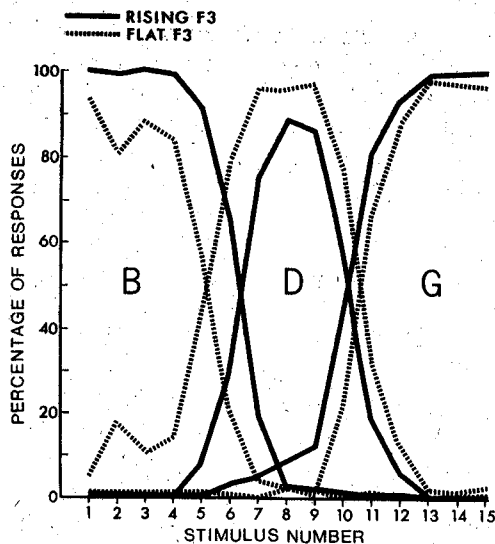

FIG. 1. Experiment I: Percentages of B, D, and G responses to binaurally presented stimuli with a rising or flat third formant (F_3).

reflected in the percentage of B responses. In combinations with stimuli 1–5, which themselves were identified as B in binaural presentation, there were, of course, almost 100% B responses. This portion of the function is of little interest. From stimulus 6 on, B responses fell below 100%, indicating the presence of phonetic conflict. However, for all stimulus combinations on the rising-F_3 continuum, the percentage of B responses remained substantially above 50%, the level of perceptual equilibrium. Thus, stimulus 1 was perceptually dominant in all these dichotic pairs.

The category goodness model predicted that the stimulus dominance function for stimulus 1 would exhibit a peak at the location of the D–G category boundary. There was indeed a clear peak in the dominance function (solid line); however, it occurred at stimulus locations 11 and 12, and thus fell somewhat to the right of the D–G boundary (located at 10.02). Stimulus 10, which was ambiguous between /dæ/ and /gæ/, was less dominated by stimulus 1 than stimuli 11 and 12, which received predominantly G responses in binaural presen-
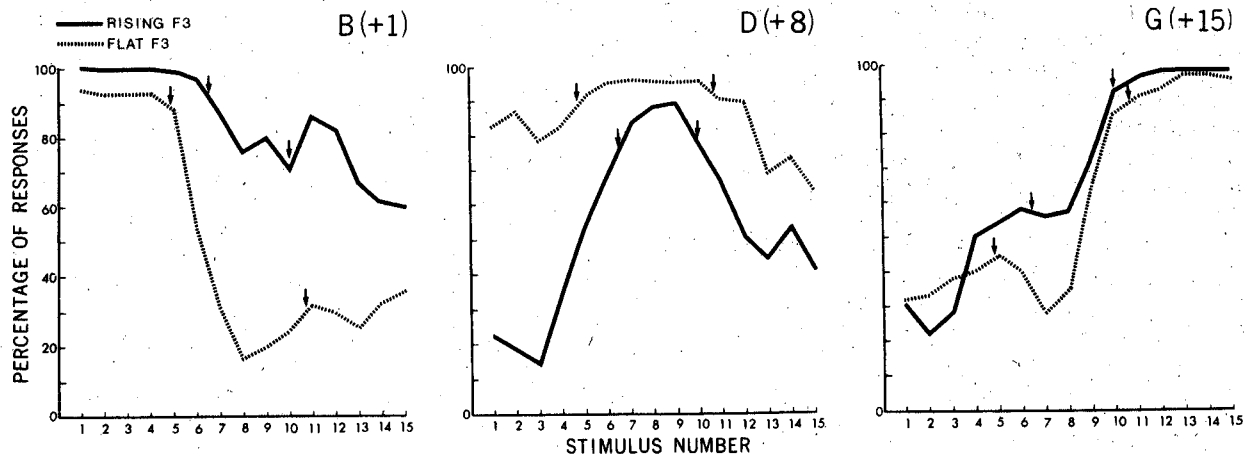

FIG. 2. Experiment I: Percentages of B, D, and G responses to stimuli dichotically fused with a constant stimulus representative of the corresponding phonetic category (+1, +8, +15).

tation. Thus, while the existence of a peak in the function would seem to support the category goodness hypothesis, its precise location does not agree with the predictions.

In the dominance function under consideration, a peak at the $B-D$ boundary cannot be discerned because of the high response ceiling to the left of the boundary. Nevertheless, the category goodness hypothesis does predict that stimuli ambiguous between /bæ/ and /dæ/ should be strongly dominated by a good /bæ/. Indeed, the pairing of stimulus 1 with the ambiguous stimulus 6 (66.7% B responses binaurally) recieved 96.8% B responses, which indicates that stimulus 6 was almost completely dominated by the less ambiguous stimulus 1. This is in agreement with the category goodness hypothesis.

Turning now to the second function in the $B(+1)$ panel of Fig. 2—that for the stimuli with a flat F_3 (dotted line)—we note first the drastic reversal in the relative dominance of stimulus 1. In all combinations with stimuli 7–15, stimulus 1 was the less dominant component, and B responses constituted a minority. The large effect of the change in F_3 on the dichotic response frequencies is surprising in view of the fact that it reduced the binaural identifiability of stimulus 1 as B by only 6.5% (from 100% to 93.5%). Nevertheless, the same change in F_3 reduced the percentage of B responses to dichotic pairings of stimulus 1 with phonetically conflicting stimuli by as much as 60%.

As in the rising-F_3 function, a peak appeared at stimulus locations 11 and 12 in the flat-F_3 function also. This peak was rather small but, because of the shift in the binaural $D-G$ boundary to the right (now located at 10.69), it was somewhat closer to the predicted location. Moreover, the most ambiguous stimulus at the $B-D$ boundary, stimulus 5 (56.9 responses binaurally), was strongly dominated by stimulus 1 (88.0% B responses, as compared with 93.5% for stimulus 1 binaurally). This lends support to the category goodness hypothesis.

Consider now the center panel in Fig. 2, which shows the dominance functions (percentage of D responses) for stimulus 8 (/dæ/) paired with all others. In the rising-F_3 set, stimulus 8 was dominated by stimuli 1–4 (/bæ/) but in approximate perceptual equilibrium with stimuli 12–15 (/gæ/). When F_3 was flat (dotted line), on the other hand, stimulus 8 dominated all other stimuli on the continuum, despite the fact that a natural /dæ/ typically has a falling F_3 transition. The change in dominance was especially pronounced in competition with stimuli 1–4, in agreement with the results shown in the $B(+1)$ panel.

Since both category boundaries are adjacent to the critical response category, D, we cannot look for distinct peaks in the $D(+8)$ functions, but stimulus 8 nevertheless should have strongly dominated ambiguous stimuli in the boundary regions. Taking the clearest example, D responses to stimulus 10 with a rising F_3 changed from 57.4% binaurally to 78.2% when fused with stimulus 8 (as compared with 88.4% to stimulus 8 binaurally). This indicates only moderate dominance of

stimulus 8 over the ambiguous stimulus 10, so that no strong support for the category goodness model can be derived from this portion of the data.

Finally, consider the $G(+15)$ dominance functions in the right-hand panel of Fig. 2. Stimulus 15 (/gæ/) with a rising F_3 tended to be dominated by /bæ/ stimuli but was in approximate equilibrium with /dæ/ stimuli (solid line). Elimination of the F_3 transition made /gæ/ weaker in competition with /dæ/ stimuli but slightly stronger in competition with /ɟæ/ stimuli (dotted line).[4] Both dominance functions were decidedly nonmonotonic in the region of the $B-D$ boundary, as predicted by the category goodness hypothesis: They exhibited broad peaks whose centers coincided approximately with the boundary locations. (The solid function does not show a peak but a plateau; however, this may be interpreted as a peak superimposed on a steep gradient.) Also, ambiguous stimuli in the $D-G$ boundary region were strongly dominated by stimulus 15. For example, stimulus 10 with a rising F_3 received 41.2% G responses binaurally, but 92.1% G responses when fused with stimulus 15.

It would be good to know whether the various peaks in the dominance functions were significant, or whether they were simply due to variability in the data. Unfortunately, there is no easy significance test, since the overall shapes of the stimulus dominance functions cannot be predicted; moreover, they varied from subject to subject. Instead of a numerical test, a qualitative summary of the individual data is presented in Table II. This summary indicates that the peaks were reasonably consistent across subjects.

C. Discussion

The results just discussed confirm Repp's (1976a) finding that stimulus dominance functions obtained by pairing a constant stimulus with all other stimuli along a place-of-articulation continuum exhibit local peaks in the vicinity of category boundaries. Such peaks were predicted by the category goodness hypothesis outlined in the Introduction. However, the peaks were not located exactly at the boundaries but typically somewhat more toward the ends of the continuum. They also failed to shift as a consequence of changes in F_3. This creates a problem for the category goodness assumption.

The possibility that the peaks in the dominance functions were determined by purely auditory factors must be given serious consideration. However, if there is an auditory explanation, it will not be a simple one. It has been hypothesized that phonetic category boundaries generally coincide with points of natural psychoacoustic discontinuity (e.g., Kuhl and Miller, 1978). However, in the case of place-of-articulation distinctions, there is as yet no clear evidence in support of this hypothesis (cf., Bailey, Summerfield, and Dorman, 1977). If some psychoacoustic discontinuities could be shown to coincide with the phoneme boundaries, then it might be argued that auditory cues are less salient in the vicinity of these boundaries because they are less salient at points of psychoacoustic discontinuity. However, there

TABLE II. Peaks in stimulus dominance functions in vicinity of distant category boundaries: qualitative summary of individual data in experiment I.

Subject	B(+1)		G(+15)	
	Rising F_3	Flat F_3	Rising F_3	Flat F_3
BHR	large (11–12)	small (10–11)	small (6)	small (5)
KH	large (11–12)	floor	none	small (4)
NK	ceiling	small (11–12)	large (4–5)	large (3–5)
SM	ceiling	small (11)	hump (6–7)	large (4–6)
WT	large (11–15)	small (12)	hump (4–7)	small (4–7)
JK	hump (11–12)	none	large (3–5)	small (5–6)

Note. Numbers in parentheses are the stimulus locations over which peaks extended. A "large" peak generally is an elevation of 20% or more. A "hump" is a peak superimposed on a steep gradient [as in the rising-F_3, $G(+15)$ function in Fig.2]. Cases where no peak could be distinguished are indicated by "ceiling," "floor," or "none," depending on the overall level of the function in the critical region; when a function was at the ceiling, a peak may have existed but could not be observed.

is no direct support even for this latter contention.

An explanation in auditory terms is complicated by the fact that the present binaural category boundaries, as well as the peaks in the dichotic stimulus dominance functions, were not symmetrically located with respect to the stimulus with a flat F_2 transition, stimulus 7 on the continuum (cf., Figs. 1 and 2): The $B-D$ boundary was closer to stimulus 7 than the $D-G$ boundary. The psychoacoustic thresholds for detecting the presence and/or direction of rising and falling F_2 transitions would be expected to be roughly equidistant from the stimulus with no transition at all. Of course, it is possible, even likely, that these thresholds critically depend on the frequencies and trajectories of other formants that are simultaneously present. For example, the rising F_1 transition in the present stimuli may have made rising F_2 transitions easier to detect than falling F_2 transitions, although this seems somewhat counterintuitive. On the other hand, the locations of perceptual boundaries for phonetic distinctions are often found to agree with the acoustic consequences of natural speech production. Therefore, the hypothesis that identification of speech sounds is somehow mediated by relationships to internal prototypes (Repp, 1976a, 1977a; Oden and Massaro, 1978) retains its plausibility. But is the dichotic competition of speech sounds governed by such relationships?

The case for the category goodness hypothesis is further weakened by the results of several recent studies, all employing two-formant synthetic stimuli in dichotic presentation, that did not obtain the peaks at category boundaries found in experiment I (Whittaker and Porter, 1976; Pompino et al., 1977; Repp, 1978c). As will be seen, the present experiment II, which employed three-formant stimuli somewhat different from those of experiment I, also failed to replicate that aspect. This accumulating evidence suggests that the stimulus dominance pattern in experiment I reflected complex, and as yet undefined, psychoacoustic properties of the specific stimuli used. Leaving the uncovering of these properties for future research, experiment II instead attempted to find support for the idea underlying the category goodness hypothesis, viz., that there exists a "multicategorical" stage of stimulus representation be-

fore dichotic information is integrated into a single phonetic percept.

II. EXPERIMENT II

One of the most reliable findings in speech perception research is the categorical perception of stop-consonant-vowel syllables varying in the formant transitions that cue place of articulation. Syllables from such a continuum invariably exhibit sharp category boundaries in a labeling task, and their discrimination is little better than predicted under the assumption that all a listener retains of a stimulus is its phonetic label (Liberman et al., 1957; Pisoni, 1971).

However, Repp (1976b) recently reported a curious result that does not fit the customary pattern. His stimuli were fused hybrid syllables composed of two different inputs to the two ears, similar to the stimuli of experiment I. He presented these fused stimuli in a discrimination task and compared the obtained performance with predictions obtained from earlier identification data for the same stimuli (Repp, 1976a). In the discrimination task, the stimulus in one ear was held constant, so that the difference to be detected occurred only in one component of the fused stimuli. Performance was predicted to be very poor, since fusion with a constant stimulus greatly reduced the phonetic distinctiveness of the stimuli. However, the obtained discriminations scores were much better than predicted and, moreover, exhibited characteristic peaks that coincided with those obtained in a control condition in which the stimuli were presented binaurally (not fused with a constant stimulus). This seemed to provide an instance of categorical perception without clearly defined categories—a paradoxical result.

This finding was especially interesting, since it could be explained by the assumption underlying the category goodness hypothesis—that the integration of information from the two ears takes place at a level intermediate between auditory and phonetic representations. At this intermediate, "multicategorical" level, a stimulus is assumed to be represented as a vector whose elements are the perceptual distances of the stimulus from the several relevant category prototypes in auditory space. Dichotic fusion of two stimuli is assumed to re-

sult in the averaging of their multicategorical vectors. If it is further assumed that the discrimination of two stimuli is based on the distance between their vectorial representations, the paradoxical results of Repp (1976b) can be explained. Obviously, the relative distances among the members of a stimulus set remain unchanged when all stimuli are fused with the same constant stimulus, i.e., when the same constant vector is added to all individual stimulus vectors; only the absolute magnitude of the distances decreases. This corresponds to a decrease in overall discriminability, without any change in the relative discriminability of different stimulus pairs, as reflected in the peaks and troughs of the discrimination function. This is precisely what Repp (1976b) obtained.

Unfortunately, there were some procedural problems in these earlier experiments. Having realized that the temporal alignment of the dichotic stimuli contained a random error that may have led to artifacts in the discrimination task, Repp (1976b) replicated the experiment, taking great pains to align the stimuli on the two tape tracks as precisely as possible, and obtained essentially the same results. However, it has since transpired that, unbeknownst to him, the specific procedure he used in the replication experiment (two-channel output of stimuli digitized at a 20-kHz sampling rate) did not function properly at the time and may have resulted in intensity and quality differences on the two tape channels. It would be difficult to explain how the quite different technical problems in the two experiments could have led to similar patterns of results that, moreover, resembled those obtained in a binaural control condition. Nevertheless, it seemed advisable to conduct another replication experiment that is free from all previous procedural problems. This was the purpose of experiment II. In addition, the experiment compared performance in dichotic and mixed (binaural) presentation conditions.

A. Method

1. Subjects

The subjects were eight paid volunteers recruited from Yale University. All had participated in at least one earlier experiment using synthetic speech and had proven to be accurate listeners.

2. Stimuli

The stimuli were eight syllables from a "place continuum" ranging from /ba/ to /da/ to /ga/, created on the OVEIIIc serial resonance synthesizer at Haskins Laboratories. All syllables were 295 ms in duration and had a constant fundamental frequency of 94 Hz. They had no release bursts and differed only in the transitions of the second and third formants (F_2 and F_3) which occupied the first 40 ms. The transition onset frequencies and steady state frequencies of F_2 and F_3 are shown in Table III. In addition, each stimulus had a transition in F_1 which rose from 285 Hz to a steady state of 771 Hz; F_4 and F_5 were hardware-fixed. All transitions were stepwise linear in 5-ms time segments.

The stimuli were digitized at 8 kHz using the Haskins

Laboratories Pulse Code Modulation (PCM) system. The onset of the first sampling period was time-locked to stimulus onset, as was the occurrence of the first pitch pulse in synthesis. Two dichotic tapes were prepared: one for identification, the other for discrimination.

The *identification tape* contained five blocks of 42 stimuli each. These 42 stimuli were a random sequence of 16 identical pairs (two presentations of each stimulus paired with itself) and 26 nonidentical pairs (all pairings of stimulus 1 with stimuli 2–8, and of stimulus 8 with stimuli 2–7, in both channel assignments). The interstimulus interval was 3 s, and blocks were separated by 6 s.

The *discrimination tape* contained four blocks of AXB triads.[5] Blocks 1 and 4 contained only identical (binaural) pairs of stimuli, i.e., each individual stimulus in a triad consisted of the same stimulus recorded on both channels. There were the same 50 AXB triads in each of these blocks, including all seven 1-step (1 versus 2, 2 versus 3, etc.) and all six 2-step (1 versus 3, 2 versus 4, etc.) stimulus comparisons in all four AXB arrangements (AAB, ABB, BAA, BBA). By mistake, two AXB triads (1-3-3 and 7-5-5) were omitted, reducing the number of triads per block from $(7+6) \times 4 = 52$ to 50. Blocks 2 and 3 constituted a single series of 208 AXB triads, divided by a pause in the middle. These 208 triads resulted from the following design: The same 52 stimulus triads as in blocks 1 and 4 (i.e., all 1-step and 2-step comparisons in all AXB arrangements, with no omissions) occurred on one channel, while a constant stimulus occurred on the other channel. The constant stimulus was either stimulus 1 or stimulus 8, and it could occur on either of the two channels, so that, all in all, there were four times as many triads as in blocks 1 or 4 ($4 \times 52 = 208$). The interstimulus intervals were 500 ms within triads and 3 s between triads.

3. Procedure

Each subject was tested in two 2-h sessions. In one of these sessions (the dichotic condition), the two channels of the tapes were directed to different ears. In the other session (the mixed condition), the two channels were electronically mixed and presented binaurally. The sequence of the dichotic and mixed conditions was counterbalanced across subjects. Since the stimuli on the two channels were exactly simultaneous and in

TABLE III. Onset frequencies and steady states of F_2 and F_3 in the stimuli used in experiment II.

Stimulus	F_2 (Hz)	F_3 (Hz)
1	859	1795
2	1037	2150
3	1224	2502
4	1404	2998
5	1588	2998
6	1770	2502
7	1770	2197
8	1770	1902
Steady states	1233	2520

phase, all stimulus pairs were perceived as single syllables originating in the center of the listener's head, both in dichotic and in mixed presentation. In each session, the identification tape was presented first and repeated once, after the tape recorder channels had been electronically reversed. The discrimination tape was also repeated once, again with channels being reversed prior to repetition.

The subjects were fully informed about the nature of the stimuli. They were instructed to respond with B, D, or G in the identification task, guessing if necessary. In the discrimination task, the response was to be A if the second stimulus was identical to the first and B if the second stimulus was identical to the third; again, guessing was required in the case of uncertainty.

The tapes were played back on an Ampex AG-500 tape recorder, and the subjects listened over Telephonics TDH-39 headsets. The electronic mixer was built at Haskins Laboratories. The amplitudes of the stimuli in the two channels were carefully equalized at a comfortable listening level, using a Hewlett-Packard voltmeter. In the mixed condition, the output was attenuated by 10 dB after mixing; this made the amplitudes approximately equal in the dichotic and mixed conditions.

B. Results

1. Identification: Identical pairs

The response to pairs of identical stimuli constituted the baseline identification data for the eight syllables. The labeling probabilities were expected to be unaffected by the mode of presentation—binaural or mixed.

Figure 3 shows that this was true. In this figure, the top panels are for the dichotic condition, and the bottom panels are for the mixed condition; the three panels in each row represent B, D, and G responses, respectively, as a function of stimulus number. The data for identical pairs are represented by the solid lines. It can be seen that the dichotic and mixed results were practically identical, and that the two category boundaries—between B and D and between D and G—were unusually sharp. There was a complete switch from B to D percepts between stimuli 3 and 4, and the change from D to G between stimuli 6 and 7 was almost as abrupt. In view of the fact that these represent the average results of eight subjects, the consistency of the labeling responses is quite remarkable; they reflect favorably on the quality of stimuli and listeners.

2. Identification: Nonidentical pairs

The labeling results for nonidentical pairs are shown as the broken lines in Fig. 3. The dashed line represents pairings of the stimuli on the abscissa with the constant stimulus 1 (+1 pairs); the dotted line represents pairings with the constant stimulus 8 (+8 pairs). Ear of presentation has been ignored in this analysis; the relevant data are presented in the Appendix.

Stimulus 1 was perceived as B in isolation. Therefore, it was expected that the labeling probabilities for +1 stimulus pairs would be biased toward B, relative to the labeling probabilities for identical pairs. The extent of the bias reflects the degree to which stimulus 1 perceptually dominated phonetically conflicting stimuli. Figure 1 shows that stimulus 1 was a weak dichotic
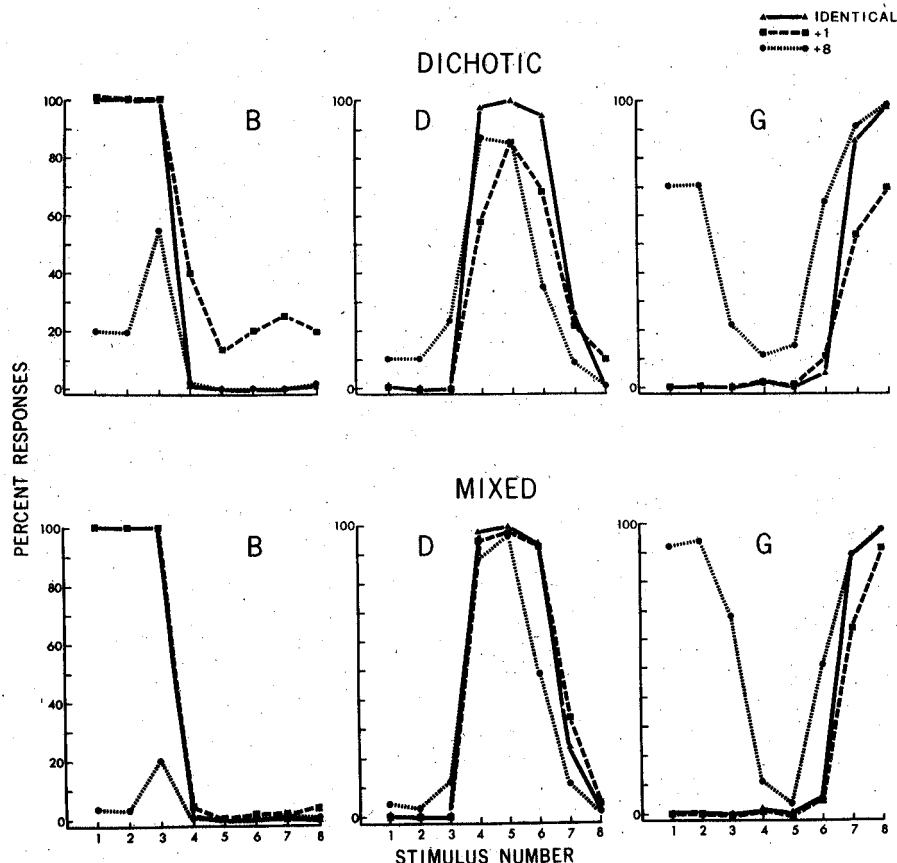

FIG. 3. Experiment II: Percentages of B, D, and G responses to identical and nonidentical stimulus pairs in dichotic and mixed presentation.

competitor: B responses to its pairings with stimuli 5–8 reached only 20%, on the average, indicating that B was strongly dominated by both D and G. Only in the pairing with stimulus 4, approximate equilibrium of responses was reached (40% B responses, 60% D responses). Mode of presentation affected the pattern of responses: In the mixed condition, B was completely dominated by D and G, so that there were practically no B responses to pairings of stimulus 1 with stimuli 4–8. This result is in marked contrast to experiment I and Repp (1976b), where B (/bæ/) tended to dominate other stimuli (/dæ/, /gæ/).

Turning now to the +8 pairs, the most relevant information is contained in the right-hand panels of Fig. 1, which show G responses. Since stimulus 8 was heard as G, it was expected to increase the number of G responses when paired with other stimuli, relative to the labeling function for these other stimuli in identical pairs. It can be seen that this was the case, but the extent of the increase in G responses varied widely depending on the nature of the competing stimulus. More specifically, G (stimulus 8) dominated B (stimuli 1 and 2), but it was dominated by D (stimuli 4 and 5). This resulted in a U-shaped function for G responses. In mixed presentation, this trend was even more pronounced than in dichotic presentation.

Clearly, D was the most dominant category in these stimuli. This can be seen best in the center panels of Fig. 1, which represent D responses. The labeling function was changed relatively little by adding either stimulus 1 or stimulus 8 to the stimuli on the abscissa, particularly in mixed presentation.[6]

The unexpectedly strong stimulus dominance effects in this experiment had the consequence of maintaining fairly distinct category boundaries even in +1 and +8 stimuli. In this respect, the present experiment was quite different from that of Repp (1976b), where stimulus dominance effects were less extreme, so that category boundaries were blurred by fusion with a constant stimulus. The precise reason for the radically different patterns of stimulus dominance relationships in the two experiments is not known at present, but there were a number of differences in the stimuli that could account for the difference. While the present experiment still provided a valid test of the question whether fused syllables are perceived categorically or not, the conditions under which this test was conducted were considerably more constrained than in the earlier experiment.

3. Discrimination: Identical pairs

The discrimination of identical pairs was expected to follow the familiar pattern of categorical perception: high performance across category boundaries and low performance within categories. Because of the unusually sharp category boundaries, this pattern was expected to be especially pronounced in the present data. Predictions were derived from the labeling probabilities using the standard formula given in Pollack and Pisoni (1971). This formula assumes that all the information the listener has available are the category labels of the stimuli, and that the stimuli in an AXB triad are categorized independently. The predictions were computed separately for each subject and then averaged.

The left-hand panels in Fig. 4 show obtained and predicted discrimination scores for identical pairs in the
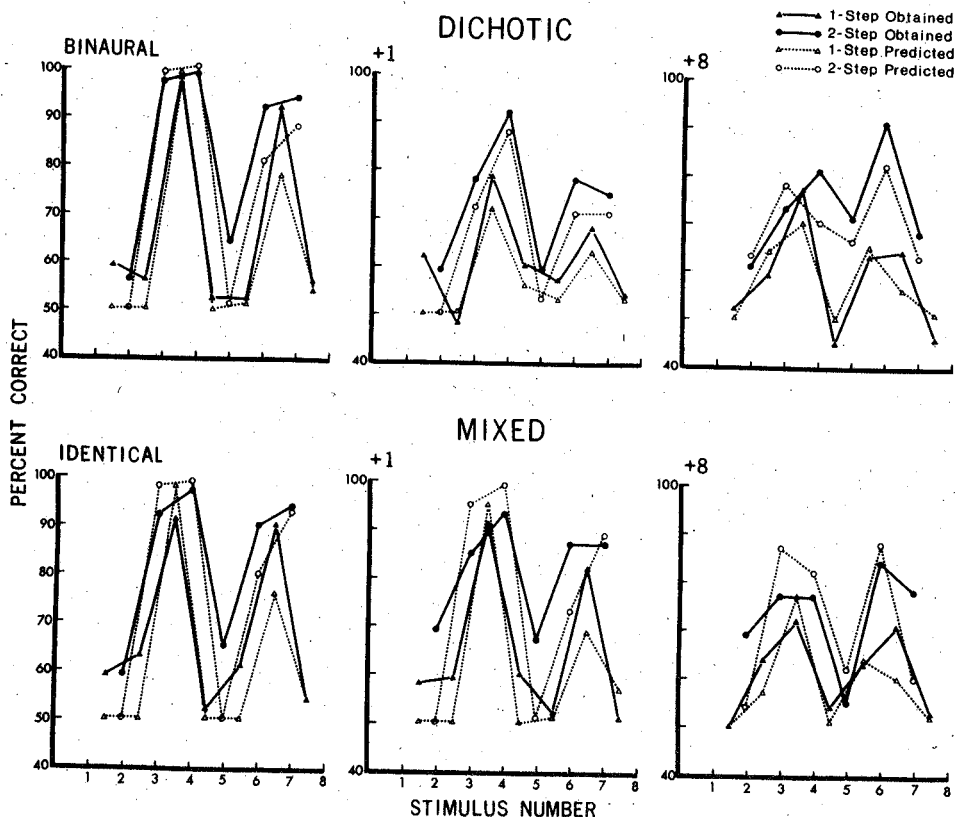

FIG. 4. Experiment II: Predicted and obtained one-step and two-step percent correct discrimination scores for identical (binaural) and non-identical stimulus pairs in dichotic and mixed presentation. Data points are plotted midway between the stimulus components to be discriminated.

dichotic and mixed conditions, respectively. In each panel, there are four functions: obtained one-step discrimination scores (triangles, solid line), obtained two-step scores (circles, solid line), predicted one-step scores (triangles, dotted line), and predicted two-step scores (circles, dotted line). It can be seen that the predicted extreme peaks and valleys in the discrimination functions were indeed obtained. The match between predicted and obtained functions was generally good, with obtained performance being somewhat higher than predicted, which is a common finding. Performance in the dichotic and mixed conditions was extremely similar, as predicted. These results indicate that the stimuli were perceived highly categorically in the absence of dichotic competition.

4. Discrimination: Nonidentical pairs

The remaining panels of Fig. 4 show the predicted and obtained discrimination results for +1 and +8 pairs in the dichotic and mixed conditions. The functions are exactly analogous to those for identical pairs, except that the stimuli to be discriminated were fused with a constant stimulus. Consider first the +1 condition. Since, as we have seen, stimulus 1 was perceptually dominated by all other stimuli and thus had only a very small effect on labeling probabilities in +1 pairs, discrimination scores in +1 pairs were also expected to be fairly similar to those for identical pairs. The data confirmed these expectations. Although discrimination scores were generally somewhat lower than for identical pairs, particularly in the dichotic condition, the pattern of results was quite similar, and the match between predicted and obtained +1 discrimination functions was quite good, particularly in the dichotic condition. Again, obtained scores exceeded predicted scores by some margin, probably reflecting an auditory memory component.

Because of the marginal effect that fusion with stimulus 1 had on performance, the +1 condition could hardly serve to test the model of dichotic competition outlined in the Introduction. However, the +8 condition provided a better opportunity to do so. As can be seen in the right-hand panels of Fig. 4, predicted and obtained scores for +8 pairs deviated from those for identical pairs, due to the larger effect that fusion with stimulus 8 had on perception. We note that, again, the fit between obtained and predicted functions was reasonable, although somewhat less convincing than with the other stimulus pairs.

According to the hypothesis that fused stimuli are perceived categorically, there should be no systematic deviations between predicted and obtained discrimination functions, except perhaps for a slight difference in overall level, due to auditory memory. On the other hand, the hypothesis, outlined in the Introduction, that discrimination is based on "multicategorical" stimulus representations predicts that the obtained +8 discrimination functions should match the pattern for identical pairs (shifted to a lower level of accuracy) more closely than the pattern predicted for +8 pairs. In order to decide this issue, we need to search for instances where the predicted functions for identical and +8 pairs have

different trends. For example, scores for the 2–4 comparison were predicted to be higher than for the 3–5 comparison in the +8 condition, whereas equal performance, or perhaps even higher performance for 3–5, was predicted in identical pairs. The results in both the dichotic and mixed +8 conditions were closer to the latter than to the former. Another, especially clear instance is the relationship between comparisons 5–7 and 6–8, for which opposite directions were predicted in identical and in +8 pairs. As can easily be seen in Fig. 2, in this case the results in the +8 condition are closer to the +8 predictions. There are three relationships between 1-step comparisons that may be similarly examined. For two of these (1–2 versus 2–3, and 4–5 versus 5–6), the +8 results go with the +8 predictions. For the third (5–6 versus 6–7), another very clear difference between identical and +8 pairs, the +8 results deviate from the +8 predictions in the direction of the results for identical pairs, but they do not nearly approach the extreme difference observed there.

In summary, the data fail to provide consistent evidence that obtained +8 scores are closer to those for identical pairs than to the predicted +8 scores. Although there are some trends in this direction, the overall evidence is too weak to reject the categorical perception hypothesis. It must be concluded that the fused stimuli in the present experiment were, in essence, perceived categorically in both dichotic and mixed modes of presentation.

C. Discussion

The main conclusion of the present experiment is that the fused syllables were categorically perceived. Whatever deviations occurred from the predicted performance pattern were not systematic enough to warrant interpretation. Thus, the listeners in the present experiment apparently relied on phonetic category labels in making their discrimination responses, regardless of whether the syllables were binaural singles or dichotic (or mixed) hybrids.

The contrasting results of Repp (1976b) had suggested that an earlier, multicategorical representation of fused dichotic stimuli can be accessed in a discrimination task. The possibility remains that such an earlier level can be utilized only when the stimuli are highly ambiguous at the categorical phonetic level. As pointed out above, this prerequisite was not sufficiently met in the present experiment. There is increasing evidence in the literature that categorical perception can be transcended by leading subjects to ignore phonetic categories and focus on auditory differences, either by long practice or by presenting only stimuli from a single phonetic category (e.g., Carney, Widin, and Viemeister, 1977; Ganong, 1977). Thus, in order to assess the nature of earlier levels of stimulus representation, phonetic categorization must somehow be prevented or deemphasized. This hypothesis might be further tested by presenting fused hybrid syllables in some of the experimental paradigms designed to reduce the role of phonetic categorization.

By indicating that the subjects did not make use of an

earlier, multicategorical stimulus code, the present results do not imply that such a level does not exist. Therefore, the failure to replicate the findings of Repp (1976b) cannot be taken as evidence against the hypothesis that information from the two ears is combined at such a multicategorical stage. One detail of the present data, however, is relevant to that hypothesis; and, unfortunately, it is not favorable. The category goodness hypothesis, which is implied by the multicategorical hypothesis, predicts that stimuli close to a phonetic category boundary should be weaker in dichotic competition than stimuli farther removed from a boundary (cf., experiment I). Figure 4 shows, however, that stimulus 3 was stronger in competition with stimulus 8 than stimuli 1 and 2 (note the peaks in the dotted functions in the left-hand panels of Fig. 3). While such a detailed result may not be sufficient by itself to reject the multicategorical hypothesis, it does add to the increasing evidence that the underlying model does not fit the detailed structure of dichotic data very well.

III. EXPERIMENT III

Experiment III took a different, and perhaps more direct, approach to testing the category goodness hypothesis of dichotic competition. In contrast to experiments I and II, the stimuli used were syllables contrasting in the *voicing* feature of the initial stop consonant (/da/–/ta/). It is known that the dichotic stimulus dominance relationship between two fused syllables contrasting in voicing can be substantially altered by changing the voice onset times (VOTs) of the competing stimuli, particularly the VOT of the voiceless stimulus (Miller, 1977; Repp, 1977a, 1978a). A /ta/ with a long VOT dominates a /da/ in the other ear more than does a /ta/ with a short VOT. This is in agreement with the category goodness hypothesis, since the prototypical voiceless stop probably has a fairly long VOT. However, the effect of variations in VOT could also be explained by interactions at the auditory level. For example, the longer aspirated portion that goes with a longer VOT may simply be more detectable or more salient in a fused stimulus composed of /da/ and /ta/. Thus, the dependence of stimulus dominance on the VOTs of the competing stimuli cannot be taken as unequivocal support for the category goodness hypothesis.

The results of recent experiments (Repp, 1979) suggested an alternative way of improving the category goodness of a /ta/. In these studies, it was shown that an increase in the amplitude of the aspiration noise in stimuli from a /da/–/ta/ continuum led to more voiceless responses and could be traded in perception for a decrease in VOT. By making all stimuli more /ta/-like, the increase in aspiration amplitude increased the category goodness of stimuli that were perceived as /ta/ to begin with. Thus, changes in the amplitude of the aspirated portion (relative to the following periodic portion) may be used to vary the dimension of category goodness of /ta/ stimuli in a dichotic experiment, in the same way that changes in VOT have been used in earlier dichotic studies (Miller, 1977; Repp, 1977a, 1978a).

There are two ways of implementing a change in aspiration amplitude. One is to lower the amplitude of the periodic portion (V) while holding the amplitude of the aspirated portion (A) constant. In the experiment of Repp (1979), the resulting reduction in the overall amplitude of the syllable (which is determined primarily by the periodic portion) had no effect on the responses. In the context of a dichotic experiment, however, the procedure just mentioned creates a situation where the overall amplitude of /ta/ will be reduced relative to that of a competing /da/. If overall amplitude differences between the competing stimuli do not affect their dominance relationship (and there was some reason to believe that this might be the case—see below), the category goodness hypothesis predicts that a /ta/ with a higher A/V ratio will be more dominant over a /da/ than a /ta/ with a lower A/V ratio, even though its overall amplitude is lower. Let us call this the strong prediction.

Of course, there is another way of changing the A/V ratio of a /ta/, viz., by increasing the amplitude of its aspirated portion, holding everything else constant. This manipulation would slightly increase the overall amplitude of the /ta/ stimulus, and the expected resulting increase in its dominance over a /da/ would not be in conflict with (but most likely larger than) a possible effect on stimulus dominance of the overall amplitude relationship between the competing dichotic stimuli. This expectation was the weaker prediction of the category goodness hypothesis.

Thus, if the amplitudes of A and V are varied orthogonally in a /ta/ that is dichotically paired with a /da/, and if amplitude relationships across ears play no role (within the limits of the experiment), then one might expect independent effects of A and V, and the subjects' responses should be a direct function of A/V ratio in the voiceless stimulus, as is the case with single stimuli presented binaurally (Repp, 1979).

Experiment III also investigated directly the role that overall amplitude relationships between the competing stimuli might play in this particular situation. For this purpose, variations in the amplitude of the /da/ stimulus were included as an additional factor. Several earlier dichotic studies have examined the role of stimulus amplitude and found it to have relatively little effect. In unfused dichotic stimuli, changes in amplitude relationships seem to play a minor role as long as one amplitude does not get too low, so that the intelligibility of the stimulus in one ear is impaired (Cullen *et al.*, 1974; Speaks and Bissonette, 1975). There is some preliminary evidence from experiments using fused stop-consonant-vowel syllables contrasting in the initial formant transitions only (i.e., place-of-articulation contrasts) that moderate attenuation of the stimulus in one ear has no perceptual consequences (Repp, 1976b). The present study is the first to ask the same question about fused syllables contrasting in VOT.

Although some earlier studies suggest that amplitude relationships are unimportant, it is dangerous to generalize from one situation to another, in view of the different psychoacoustic situations represented by different dichotic experiments. In contrast to syllables that do not fuse, and in contrast to dichotic place contrasts

that fuse perfectly, dichotic voicing contrasts are "partially fused," given that they differ only in VOT (Repp, 1978a). The initial aspirated portion of the voiceless stimulus does not fuse with the initial (voiced) portion of the voiced stimulus, but the remaining, identical periodic portions in the two stimuli fuse perfectly and create a single auditory image localized between the two ears. Earlier experiments (Repp, 1977a, 1978a) have shown that listeners can perceptually integrate all these stimulus components into a single phonetic percept; yet it is true that a careful listener can easily determine the ear in which the aspiration noise occurs. This—from a psychoacoustic viewpoint—quite unique situation justifies a separate inquiry into the effects of amplitude relationships on stimulus dominance. Needless to say, such amplitude effects, if found, could not be explained by the category goodness hypothesis, but rather would have to be ascribed to psychoacoustic factors.

A. Method

1. Subjects

Eight subjects participated. They included the author, a research assistant, and six paid volunteers (Yale undergraduates) who had participated in previous experiments using synthetic syllables and had proven to be reliable listeners.

2. Stimuli

The stimuli were generated with the OVEIIIc serial resonance synthesizer at Haskins Laboratories. They were stop-consonant-vowel syllables perceived as either /da/ or /ta/. Their total duration was 300 ms. Fundamental frequency was constant at 125 Hz over the first 84 ms and then fell linearly to 90 Hz at offset. The initial formant transitions were stepwise-linear and 48 ms in duration; F_1 rose from 285 to 771 Hz, F_2 fell from 1543 to 1233 Hz, and F_3 fell from 3019 to 2520 Hz. The duration of the synthesis time frames was 4 ms.

The change from /da/ to /ta/ was created by replacing periodic excitation with noise and simultaneously increasing the bandwidth of F_1 to its maximum (thereby essentially eliminating F_1). The amplitude of the aspiration noise was about 20 dB below that of the following periodic portion, as determined by later measurements of the synthesizer output. The periodic source was turned on 8 ms (one pitch period) before voicing onset but kept at a minimal amplitude. This procedure insured that the second pitch pulse, which marked the true onset of voicing, had full amplitude. The stimuli had no special release bursts at onset. There were two basic stimuli: a /da/ with a VOT of 0 (i.e., no aspiration at all) and a /ta/ with a VOT of 44 ms.[7]

The two basic stimuli were digitized at 10 kHz using the Haskins Laboratories pulse code modulation system. From the digitized waveforms, eight additional versions of /ta/ were constructed by independently amplifying or attenuating the aspirated and periodic portions of the original stimulus. Changes in amplitude in either stimulus portion were achieved by means of a computer instruction after placing a cursor at the onset of the first

true pitch pulse. Relative amplitudes of the aspirated portion of either −6, 0, or +6 dB (relative to the original stimulus) were orthogonally combined with relative amplitudes of the periodic portion of either −6, 0, or +6 dB (relative to the original stimulus). Thus, the stimulus ensemble included both a 12-dB range in absolute stimulus amplitude (from −6/−6 to +6/+6; the slash symbolizes the partition into aspirated and periodic stimulus portions), and a 24-dB range in the relative amplitudes of aspirated and periodic portions in /ta/ stimuli (from −6/+6 at one extreme to +6/−6 at the other). Given a true amplitude ratio between periodic and nonperiodic portions of about 20 dB in the original (0/0) /ta/ stimuli, the total range extended from 8 dB (+6/−6) to 32 dB (−6/+6).

The independent amplitude variations in the aspirated and periodic portions of /ta/ will be denoted by A and V_T, respectively, to distinguish the latter from the variations in the amplitude of the (all-periodic) /da/, denoted by V_D. In the experiment, all tokens of /da/ were paired dichotically with all tokens of /ta/. Thus, the experiment had a four-way factorial design: A (−6, 0, +6 dB), V_T (−6, 0, +6 dB), V_D (−6, 0, +6 dB), and ear of presentation (/ta/ in left or right ear). The orthogonal combination of all four factors led to 54 stimuli that were replicated once and recorded in four different random sequences of 108, with ISIs of 3 s. The onsets of the dichotic stimuli were perfectly simultaneous, and the identical periodic portions were always exactly in phase.

3. Procedure

The stimulus tape was presented twice, in separate sessions. Tape recorder channels were reversed electronically between sessions for counterbalancing purposes. All in all, each subject listened to eight blocks of 108 stimuli. The task was to identify each fused stimulus as beginning with a D or a T, guessing if necessary.

B. Results

A four-way analysis of variance was conducted on the response frequencies. There were four highly significant effects; no other effect even approached significance. Since there was no significant effect involving ear of presentation, the results were collapsed over this factor, yielding 32 responses per subject per cell in the remaining three-factor design. The effects of the three amplitude factors, V_D, A, and V_T, are graphically shown in Fig. 5.

Each panel in Fig. 5 shows the percentage of D responses as a function of A. In each panel, V_T is the parameter of the three functions (solid lines), whereas V_D increases across panels. The first effect to be seen is a general increase in D responses as V_D increased, comparing data points across the three panels, $F(2,14) = 11.5, p < 0.01$. Thus, contrary to expectations based on earlier studies, there was a clear effect of increasing the amplitude of one stimulus (/da/) on its relative perceptual dominance.

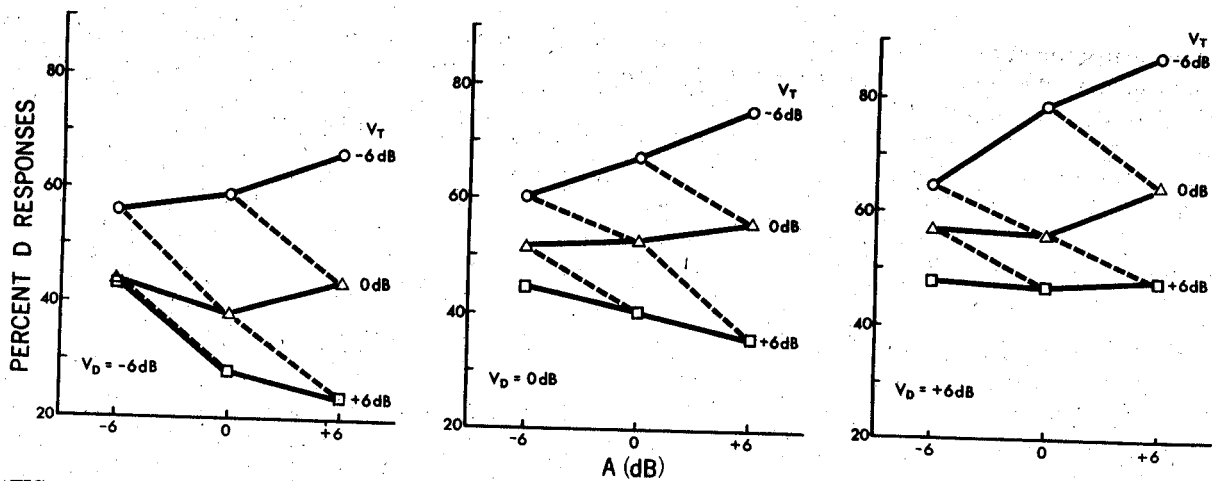The second effect can be seen in the large displace-

FIG. 5. Experiment III: Dichotic voicing contrasts—percentage of D responses as a function of A, V_T, and V_D. The dashed lines connect points of equal A/V_T ratio.

ments of the three solid functions within each panel. This reflects a main effect of V_T, the amplitude of the periodic portion of /ta/: D responses increased as V_T decreased, $F(2,14) = 9.7$, $p < 0.01$. Obviously, this effect contradicts the strong prediction of the category goodness hypothesis: D responses should have decreased (and T responses increased) as V_T decreased, thus increasing the A/V_T ratio and the category goodness of /ta/ stimuli. The effect actually obtained indicates that not category goodness but overall amplitude was the decisive factor in dichotic competition.

The third effect displayed in Fig. 5 is the fan-shaped pattern of the solid functions in each panel. It represents an A by V_T interaction, $F(4,28) = 4.8$, $p < 0.01$. When in dichotic competition with a /da/, changes in V_T had a larger effect when A was high then when it was low. Expressed differently, A—the amplitude of the aspirated portion—had different effects depending on the level of V_T, the amplitude of the periodic portion in /ta/. In fact, an increase in A had the predicted effect —a reduction in D responses—only when V_T was high (+ 6 dB). When V_T was low (−6 dB), the effect of A was inverted, higher amplitudes leading to more D (fewer T) responses! Together with the fact that A had no significant main effect, this result refutes even the weaker prediction of the category goodness hypothesis and calls for explanation. (See Sec. III C.)

The interaction just described suggests that, in contrast to the perception of single (binaural) stimuli (Repp, 1979), a constant A/V_T ratio in the /ta/ syllables did not lead to a constant perceptual result. This is confirmed by inspecting the dashed lines in Fig. 5, which connect points of constant A/V_T ratio. These lines have uniformly negative slopes, indicating that D responses decreased as the overall amplitude of the /ta/ increased, holding A/V_T ratio constant.

The fourth effect shown in Fig. 5 is represented by a rotation of the fan-shaped pattern (due to the A by V_T interaction) from a downward orientation in the left-hand panel to an upward orientation in the right-hand panel. This rotation reflects a V_D by A interaction, $F(4,28) = 5.2$, $p < 0.01$. Thus, the effect of A depended

on both V_D and V_T, being in the expected direction when V_D was low and V_T high, negligible when V_D equaled V_T, and in the opposite direction when V_D was high and V_T was low. This suggests that the subjects' responses were essentially a function of A and V_D/V_T ratio. A comparison of those solid functions representing the same V_D/V_T ratio across different panels in Fig. 5 suggests that the direction of the A effect was indeed similar in conditions of equal V_D/V_T ratio. There was a small but consistent additional effect of absolute amplitude level: An equal increase in the amplitudes of both competing stimuli (hence, of V_D, A, and V_T simultaneously) led to a slight increase in D responses.

C. Discussion

The results of experiment III clearly refute the category goodness hypothesis of dichotic competition. There was absolutely no evidence in favor of it, suggesting that effects of category goodness were truly absent, not just overcome by more powerful effects of overall amplitude. Although this negative result has been obtained in a situation with very specific psychoacoustic characteristics (partially fused syllables), it certainly raises grave doubts about whether *any* dichotic stimulus dominance effects can be explained by relative category goodness. Findings previously thought to support this model may ultimately be explainable in psychoacoustic terms.

The key to understanding the complex pattern of results in experiment III presumably lies in the unique psychoacoustic properties of the dichotic stimuli employed. Being partially fused, they consisted of a brief unfused part (aspiration noise in one ear, periodic sound in the other) followed by a perfectly fused longer periodic portion. Changes in the amplitude relationships between the two stimuli (and hence, between the two identical periodic portions) led to changes in the subjective localization of the fused image. It seems that these localization shifts can explain the perceptual findings.

Consider the case where the amplitude of /da/ is low-

er than that of the periodic portion of /ta/ ($V_D < V_T$). In this case, the fused periodic portion will be localized toward the side where the /ta/ occurred and, thus, where the aspiration noise is heard (since it does not fuse with the initial voiced portion of the /da/ in the other ear). In this case, listeners seem to perceptually integrate the aspiration noise with the following periodic portion and give a majority of /ta/ responses. Moreover, if the amplitude of the aspiration (A) is increased, holding V_D and V_T constant, it has the desired effect of further increasing /ta/ responses. Consider now the case where $V_D > V_T$. Here, the fused periodic portion will be localized toward the side where the /da/ occurred, away from the aspiration noise. In this case, listeners may find it difficult to integrate the noise with the periodic portion; therefore, there are few T responses. (This interpretation is reasonable, since informal observations suggest that a noise portion in one ear followed by a periodic portion in the other ear, i.e., a single /ta/ split between the two ears, is generally not perceived as /ta/.) When A is increased in this situation, it has the paradoxical effect of decreasing T responses. Given that the noise is already "streaming" away from the periodic portion, an increase in its amplitude probably further increases the perceptual dissociation of the two stimulus components (cf., Dannenbring and Bregman, 1976). The case of $V_D = V_T$ falls between these two extremes. Here, the fused periodic portion is localized in the midline, and the opposed effects of changes in A on the voicing decision and on the perceptual dissociation of noise and periodic portions seem to cancel out, since A has little effect on the responses.

Thus, the results can be explained by taking account of the relative positions of the stimulus portions in subjective space. The four principal effects in the data described earlier (V_D, V_T, $V_D \times A$, $V_T \times A$) are really only two: There is an effect of V_D/V_T ratio (that is, of closeness in auditory space of the aspirated portion and the fused periodic portion) and an interaction between V_D/V_T ratio and A (that is, the effect of A depends on whether the aspiration can be integrated with the fused periodic portion). The small additional effect of overall stimulus amplitude (more D responses when amplitude was uniformly increased) probably reflects a certain amount of interference of the initial periodic portion of /da/ with the perception of the simultaneous aspiration noise in the other ear, and this interference increased at higher stimulus intensities. The present data then do not really contradict earlier dichotic studies that found no amplitude effects, if the hypothesis is accepted that amplitude effects in partially fused voicing contrasts are mediated by changes in the relative localizations and relative auditory coherence of stimulus portions, due to the unique psychoacoustic properties of these stimuli. Rather, the results of experiment III provide an interesting example of how spatial separation can lead to perceptual dissociation (cf., also Axelrod, Guzy, and Diamond, 1968; Huggins, 1964, 1974, 1976; Rand, 1974). Thus, they demonstrate an often neglected problem: The dichotic presentation of two speech stimuli may lead to complex and possibly quite unique binaural

interactions that are likely to constitute major determinants of ear and stimulus dominance effects.

IV. SUMMARY AND CONCLUSION

Three detailed parametric studies of dichotic stimulus competition were conducted, designed to test the hypothesis that dichotic stimulus dominance relationships reflect the combination of information at a stage intermediate between auditory and (categorical) phonetic processing. Experiment I examined the prediction that the category goodness of a speech stimulus is reflected in its degree of dominance over other stimuli in dichotic competition. While some aspects of the results seemed to confirm the prediction, there were problems with interpreting the detailed pattern of results. The possibility was raised that the observed variations in dichotic stimulus dominance had a psychoacoustic basis, but the nature of these psychoacoustic factors could not be pinned down. Experiment II followed up some earlier results suggesting that the discrimination of fused dichotic syllables might be based upon stimulus representations at a stage between auditory processing and phonetic categorization, the stage that also gives rise to the hypothesized category goodness effect. However, the results of experiment II indicated that discrimination was based on phonetic categories (i.e., it followed the characteristic pattern of categorical perception) and thus failed to replicate the earlier result. While the results of experiment II have no bearing on the reality of the hypothetical intermediate processing stage—the subjects may have chosen, for whatever reason, not to access the stimulus representations at that stage—they certainly fail to provide positive evidence. Experiment III returned to an examination of the category goodness hypothesis, and using a method different from that of experiment I, it led to a clear rejection of the hypothesis. Moreover, a plausible psychoacoustic explanation of the perceptual results could be provided. Taken together, the results of the three experiments weigh heavily against the model of dichotic stimulus competition proposed by Repp (1976a, 1977a). They suggest that, rather, dichotic stimulus dominance relationships between speech sounds can be explained in terms of auditory stimulus properties and binaural psychoacoustic effects. Future research will have to be directed at defining and further investigating these auditory factors.

ACKNOWLEDGMENTS

APPENDIX: EAR DOMINANCE EFFECTS

Although ear dominance effects were not relevant to the issue addressed by the present studies, they did occur (as in nearly every dichotic experiment) and deserve a brief discussion. Repp (1977b) has argued that

fused syllables offer certain methodological advantages in measuring lateral asymmetries. The present experiments add to the—still very small—data base concerning ear dominance for fused syllables, and, in addition, each experiment answers some additional questions about the behavior of these laterality effects with respect to certain manipulations of stimuli or tasks.

1. Experiment I

A significant but small average right-ear dominance effect was obtained by Repp (1976a) with stimuli very similar to those of experiment I. It seemed important to establish that the present stimuli yield further evidence of significant (right-) ear dominance effects, that these effects are not sensitive to changes in F_3, and that they do not vary too much across individual stimulus pairs. Positive evidence on all three counts is a prerequisite for fused dichotic syllables to be methodologically useful in assessing hemispheric dominance for speech perception.

The ear dominance effects for the six individual subjects are shown in Table AI, separately for the two stimulus sets differing in the third formant. Two ear dominance indices are reported, e' and e. (For a detailed discussion of these, see Repp, 1977b.) Their values are generally very similar. While e' is preferable on theoretical grounds, only the e index can easily be tested for significance; therefore, both are reported here. The significance test is based on a weighted standard error derived from the variability of ear asymmetries across individual stimulus pairs. Both indices take stimulus dominance effects into account; the e' index, especially, is intended to represent an unbiased estimate of ear dominance. Its value, like that of the e index, ranges from -1 (perfect left-ear dominance) to $+1$ (perfect right-ear dominance) and represents the linearly scaled intercept of a bilinear ROC (isolaterality) function with the negative diagonal of the unit square. (See Repp, 1977b.)

The average ear dominance coefficient, although based on only six subjects, was similar to that reported by Repp (1976a). Table AI shows that three subjects (including one left-hander) were significantly right-ear dominant, while the other three subjects (including BHR, the author) showed no significant ear asymmetry. The ear dominance coefficients for the two stimulus sets

differing in F_3 were generally very similar, the correlation being $+0.91$ for e' and $+0.96$ for e. Thus, the change in F_3 did not affect the degree of ear dominance. The variability across individual stimulus pairs was considerable, however. For example, subject NK's right-ear advantage (the largest of all) was based on 32 stimulus pairs, 11 of which showed no ear dominance or left-ear dominance. (Many other stimulus pairs received only a single type of response and therefore did not enter into the estimate of e'.) Of those stimulus pairs that contributed most to the estimate of NK's ear dominance (because of their small stimulus dominance effects), two out of eleven still showed left-ear dominance. It cannot be determined from the present data whether this was merely due to large variability, or whether it represented a genuine difference in ear dominance among individual stimulus pairs.

The conclusion from these results is that fused dichotic syllables (place-of-articulation contrasts) do tend to lead to right-ear dominance, although these effects are smaller than those observed with dichotic voicing contrasts (Repp, 1977a, 1978a; see also experiment III below). Ear dominance seems to be highly variable, so that a large number of trials is needed to obtain a precise estimate. Changes in F_3 do not affect the ear asymmetry.

2. Experiment II

a. Identification

Experiment II employed perfectly fused stimuli similar to those of experiment I but of better quality. Ear dominance coefficients for individual subjects are shown in the fourth and fifth columns of Table AI. The table shows that four subjects exhibited large and significant asymmetries in favor of the right ear, one subject showed significant left-ear dominance, and the remaining three subjects showed no significant effects. This distribution resembles that found in earlier studies using fused syllables (Repp, 1976a; the present experiment I), although the average right-ear dominance is somewhat larger here. The present results include some of the largest ear dominance effects ever observed with fused syllables, thus providing clear evidence that perfect fusion does not prevent lateral asymmetries. To which extent these asymmetries actually reflect hemispheric dominance for speech perception is a question that the present experiment cannot answer,

TABLE AI. Experiment I: ear dominance coefficients.

Subject	Sex	Handedness	Rising F_3		Flat F_3	
			e'	e	e'	e
NK	M	R	+0.28	+0.25[b]	+0.30	+0.29[b]
WT	M	L	+0.22	+0.21[b]	+0.20	+0.17[b]
SM	F	R	+0.13	+0.15[a]	+0.17	+0.17[b]
BHR	M	R	+0.04	+0.01	−0.01	+0.03
KH	F	R	+0.05	0.00	−0.11	−0.08
JK	M	R	−0.08	−0.06	−0.07	−0.06
Average			+0.11	+0.08	+0.08	+0.09

[a] $p < 0.05$.
[b] $p < 0.001$.

TABLE AII. Experiment II: ear dominance coefficients.

Subject	Sex	Handedness	Identification		Discrimination
			e'	e	
SA	F	R	+0.59	+0.57 [c]	−0.01
JK	M	R	+0.67	+0.39 [a]	+0.14
IS	F	R	+0.63	+0.40 [b]	+0.13
SE	M	L	+0.23	+0.28 [c]	−0.03
JG	M	R	0.00	+0.02	−0.15
JW	M	R	−0.10	0.00	0.00
MC	F	R	−0.05	−0.14	+0.16
JM	F	R	−0.38	−0.32 [a]	−0.40
Average			+0.20	+0.15	−0.02

[a] $p < 0.05$.
[b] $p < 0.01$.
[c] $p < 0.001$.

but the higher incidence of right-ear dominance suggests the involvement of speech-specific mechanisms. One disturbing observation is that subject JK, who also participated in experiment I, showed a large right-ear advantage here but no ear asymmetry there. The cause of this discrepancy is not known.

b. Discrimination

Ear dominance coefficients for the discrimination task are reported in column 6 of Table AII, where they may be compared with the coefficients obtained in the identification task. The coefficients computed from the discrimination scores are not strictly equivalent with e' and e in the identification task, although they too range from -1 to $+1$. The coefficient used is $(R*-L*)/(R*+L*)$, where $R* = 2(R - 50)$ and $L* = 2(L - 50)$, and R and L are the percent-correct scores for the two ears. Thus, it incorporates a crude correction for guessing, but, unlike e' and e, it does not properly take into account the variability between individual stimulus pairs and therefore may underestimate the size of the ear dominance effects (cf., Repp, 1977b). The statistical significance of these coefficients could not be assessed.

The coefficients for the discrimination task correlated $+0.66$ ($p < 0.05$) with e' and $+0.55$ ($p < 0.10$) with e for the identification task, indicating some consistency in individual ear asymmetries across different tasks. What is more striking, however, is that there was no longer any overall tendency toward right-ear dominance in the discrimination task. Two of the subjects (SA, SE) who had been strongly right-ear dominant in identification lost this asymmetry completely in discrimination. This indicates that even ear asymmetries for fused syllables may be sensitive to task characteristics (cf., Haggard, 1976). The discrimination task may have induced a less clearly linguistic mode of processing, thus eliminating any left-hemisphere superiority.

3. Experiment III

Repp (1977a, 1978b) reported unusually strong ear dominance effects for fused dichotic voicing contrasts, almost invariably in favor of the right ear. The results of experiment III, in contrast to these earlier results, showed no significant overall ear advantage, although

there was a small average effect in favor of the right ear. However, individual ear dominance effects were pronounced, as shown in Table AIII.

All eight subjects showed highly significant ear dominance effects. The largest right-ear advantage was shown by the author (BHR), in agreement with many previous results; for him, the right ear was almost completely dominant. (His data nevertheless provided information about stimulus dominance: A strongly dominant stimulus in the left ear often overcame the strong right-ear dominance.) This finding may be contrasted with the absence of any ear asymmetry for this listener in experiment I (Table AI). The next-largest right-ear advantage was obtained for a familial left-hander—a type of subject who typically does not show large right-ear advantages (Hardyck and Petrinovich, 1977). There were three additional significant right-ear advantages. Of the remaining three subjects, one—a nonfamilial left-hander—showed a small left-ear advantage; the other two subjects were right handers with large left-ear advantages—a rather puzzling result. Thus, while the present study confirms the proclivity of dichotic voicing contrasts to lead to pronounced individual ear dominance effects—more so than dichotic place contrasts—it also increases the urgency of the question whether these effects reflect hemispheric dominance for speech, or perhaps some other kind of lateral asymmetry in auditory processing. The fragmentary but nevertheless convincing evidence for substantial differ-

TABLE AIII. Experiment III: ear dominance coefficients.

Subject	Sex	Handedness	e'	e
BHR	M	R	+0.93	+0.92 [b]
SB	F	L(familial)	+0.72	+0.65 [b]
DF	F	R	+0.43	+0.37 [b]
DK	F	R	+0.29	+0.27 [b]
JK	M	R	+0.16	+0.20 [b]
AM	M	L(nonfam.)	−0.11	−0.13 [a]
DW	F	R	−0.50	−0.45 [b]
MB	F	R	−0.66	−0.63 [b]
Average			+0.16	+0.15

[a] $p < 0.01$.
[b] $p < 0.001$.

ences in ear advantages between different stimuli and tasks suggests that, as fas as our insight into the causes of these ear asymmetries are concerned, we are still groping in the dark.

[1] In earlier publications, this hypothesis has sometimes been called the "prototype model" (Repp, 1976a, 1977a) or the "prototype matching hypothesis" (Repp, 1978b). However, "category goodness hypothesis" is to be preferred because it denotes the specific application of the prototype idea to dichotic stimulus competition. Rejection of the category goodness hypothesis does not imply rejection of the more general prototype model (Oden and Massaro, 1978); indeed, such a model may still provide a good description of the identification process for speech stimuli, including those resulting from the fusion of dichotic stimuli at a relatively early level in processing.

[2] At the time the experimental tape was recorded, the digital sampling procedure led to a random error in dichotic stimulus alignment of up to one sampling period (0.125 ms), which was considered insignificant.

[3] Although the basic effect of the F_3 transition on the perceived place of articulation confirmed the earlier results of Harris *et al*. (1958) and Hoffman (1958), it was much smaller than in these earlier studies (though adequate for the purpose of the present experiment). Harris *et al*. and Hoffman obtained hardly any D responses to stimuli with a rising F_3 transition less steep than the present one. The reason for this discrepancy presumably lies in the relative amplitudes of F_3 and F_2. Harris *et al*. and Hoffman, who constructed their stimuli on the pattern playback, did not report any formant amplitudes, but a cautious estimate suggests that F_3 was 6–10 dB below F_2 during the transitional portion. In the present stimuli, on the other hand, F_3 was at least 12 dB below F_2, which may explain the difference in results. Clearly, the relative amplitude of a formant determines its salience as a perceptual cue.

[4] One apparent contradiction in the data may be noted by comparing the results for the stimulus pair $1+15$ in the $B(+1)$ panel of Fig. 2 with those for the same pair in the $G(+15)$ panel. There was a large effect of the change in F_3 on B responses, but no effect at all on G responses. The solution lies in the differential occurrence of D responses. D responses to dichotic combinations of stimuli heard as B and G, respectively, in isolation, have been ascribed to "psychoacoustic fusion" by Cutting (1976). These responses were infrequent when F_3 was rising, but quite frequent when F_3 was flat. Of course, this is in agreement with the general increase in D responses to stimuli with a flat F_3. Thus, the stimulus pair $1+15$ received 9.7% D responses when F_3 war rising, but 32.2% D responses when F_3 was flat. By comparison, the percentages of D responses to stimuli 1 and 15 in isolation were 0 and 0.5, respectively, when F_3 was rising, and 5.1 and 2.8, respectively, when F_3 was flat. Thus, the effect of F_3 on the frequency of psychoacoustic fusion responses, like its effect on stimulus dominance, was much larger than its effect on binaural identification scores. This reflects the increased perceptual weight of F_3 in the presence of conflicting F_2 transitions.

[5] The AXB procedure differs from the more commonly used ABX paradigm in that the first and the third stimulus are always different from each other, while the second stimulus is identical with either the first or the third. The AXB paradigm was chosen to prevent the listener strategy of attempting to compare the first with the third stimulus— a strategy that is likely to be ineffective because the auditory traces of the first timuluus may be lost by the time the third arrives. Thus, the AXB paradigm avoids the memory limitations that may reduce performance in the ABX paradigm (Pisoni and Lazarus, 1974). Although there is, at

present, no empirical evidence that the AXB paradigm is superior to the ABX paradigm, there is no reason to believe that it would be inferior. In fact, the AXB procedure may be considered a condensed version of the 4IAX paradigm, which has been shown to lead to higher discrimination performance than the ABX paradigm (Pisoni, 1973; Pisoni and Lazarus, 1974).

[6] D responses to stimuli 1–3 paired with stimulus 8, and to stimuli 7–8 paired with stimulus 1 represent "psychoacoustic fusions" (Cutting, 1976): The subjects heard D although the components stimuli were heard as B and G, respectively, in identical pairs. Psychoacoustic fusions occurred with about equal frequency in the dichotic and mixed conditions, and they were not particularly common. Their frequency tended to increase as one of the component stimuli moved closer toward the D category; this confirms earlier results by Cutting (1976) and Repp (1976a).

[7] Actually, there were two different /ta/ stimuli, with intended VOTs of 44 and 56 ms. However, later examination of the stimulus specifications revealed an embarrassing error in stimulus synthesis: /ta/ stimuli with a VOT of 56 ms in truth had 12 ms of voicing at onset, followed by 44 of aspiration. These bizarre stimuli, though impossible in articulatory terms, did not sound anomalous and, on the average, turned out to be perceptually equivalent to stimuli with a VOT of 44 ms (which had been properly synthesized). There was no significant effect of the VOT factor and only one interaction involving VOT which did not imply a qualitative change in results. Therefore, all results were collapsed over the intended VOT difference, which is ignored in the present description of method and results. Exclusion of the results deriving from the bizarre stimuli would not have significantly changed the pattern of results reported here.

Axelrod, S., Guzy, L. T., and Diamond, I. T. (1968). "Perceived rate of monotic and dichotically alternating clicks," J. Acoust. Soc. Am. 43, 51–55.

Bailey, P. J., Summerfield, A. Q., and Dorman, M. F. (1977). "On the identification of sine-wave analogues of certain speech sounds," Haskins SR-51/52, 1–27.

Berlin, C. I., Lowe-Bell, S. S., Cullen, J. K., Jr., Thompson, C. L., and Loovis, C. F. (1973). "Dichotic speech perception: An interpretation of right-ear advantage and temporal offset effects," J. Acoust. Soc. Am. 53, 699–709.

Carney, A. E., Widin, G. P., and Viemeister, N. F. (1977). "Noncategorical perception of stop consonants differing in VOT," J. Acoust. Soc. Am. 62, 961–970.

Cullen, J. K., Jr., Thompson, C. L., Hughes, L. F., Berlin, C. I., and Samson, D. S. (1974). "The effects of varied acoustic parameters on performance in dichotic speech perception tasks," Brain Lang. 1, 307–322.

Cutting, J. E. (1976). "Auditory and linguistic processes in speech perception: Inferences from six fusions in dichotic listening," Psychol. Rev. 83, 114–140.

Dannenbring, G. L., and Bregman, A. S. (1976). "Stream segregation and the illusion of overlap," J. Exp Psychol. (HPP) 2, 544–555.

Ganong, W. F. III. (1977). "Selective adaptation and speech perception," Cambridge Ph.D. dissertation (Massachusetts Institute of Technology, Cambridge).

Haggard, M. P. (1976) . "Dichotic listening," Speech Perception, Series 2, No. 5, 40–56 (Belfast, N.I.: Queen's University, Dept. of Psychology).

Halwes, T. G. (1969). "Effects of dichotic fusion on the perception of speech," Ph.D. dissertation (University of Minnesota).

Hardyck, C., and Petrinovich, L. F. (1977). "Left-handedness," Psych. Bull. 84, 385–404.

Harris, K. S., Hoffman, H. S., Liberman, A. M., Delattre, P. C., and Cooper F. S. (1958). "Effect of third-formant transitions on the perception of voiced stop consonants," J. Acoust. Soc. Am. 30, 122–126.

Hoffman, H. S. (1958). "Study of some cues in the perception of the voiced stop consonants," J. Acoust. Soc. Am. 30, 1035–1041.

Huggins, A. W. F. (1964). "Distortion of the temporal pattern of speech: Interruption and alternation," J. Acoust. Soc. Am. 36, 1055–1064.

Huggins, A. W. F. (1974). "On perceptual integration of dichotically alternated pulse trains," J. Acoust. Soc. Am. 56, 939–943.

Huggins, A. W. F. (1976). "Is periodicity detection central?" J. Acoust. Soc. Am. Suppl. 1 59, S52(A).

Kuhl, P. K., and Miller, J. D. (1978). "Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli," J. Acoust. Soc. Am. 63, 905–917.

Liberman, A. M., Harris, K. S., Hoffman, H. S., and Griffith, B. C. (1957). "The discrimination of speech sounds within and across phoneme boundaries," J. Exp. Psychol. 54, 358–368.

Miller, J. L. (1977). "Properties of feature detectors for VOT: The voiceless channel of analysis," J. Acoust. Soc. Am. 62, 641–648.

Oden, G. C. (1978). "Integration of place and voicing information in the identification of synthetic stop consonants," J. Phonetics 6, 83–93.

Oden, G. C., and Massaro, D. W. (1978). "Integration of featural information in speech perception," Psychol. Rev. 85, 172–191.

Pisoni, D. B. (1971). "On the nature of categorical perception of speech sounds," Ph.D. dissertation, University of Michigan.

Pisoni, D. B. (1973). "Auditory and phonetic memory codes in the discrimination of consonants and vowels," Percept. Psychophys. 13, 253–260.

Pisoni, D. E., and Lazarus, J. H. (1974). "Categorical and noncategorical modes of speech perception along the voicing continuum," J. Acoust. Soc. Am. 55, 328–333.

Pollack, I., and Pisoni, D. B. (1971). "On the comparison between identification and discrimination tests in speech perception," Psychonom. Sci. 24, 299–300.

Pompino, B., Rilhac-Sutter, M., Simon, A., and Sommer, R. (1977). "Auditorische Faktoren der Gewichtung bei psychoakustischer Fusion," Forschungsberichte des Instituts fuer Phonetik und sprachliche Kommunikation 8, 97–120 (University of Munich).

Rand, T. C. (1974). "Dichotic release from masking for speech," J. Acoust. Soc. Am. 55, 678–680.

Repp, B. H. (1976a). "Identification of dichotic fusions," J. Acoust. Soc. Am. 60, 456–469.

Repp, B. H. (1976b). "Discrimination of dichotic fusions," Haskins SR-45/46, 123–139.

Repp, B. H. (1977a). "Dichotic competition of speech sounds: The role of acoustic stimulus structure," J. Exp. Psychol. (HPP) 3, 53–70.

Repp, B. H. (1977b). "Measuring laterality effects in dichotic listening," J. Acoust. Soc. Am. 62, 720–737.

Repp, B. H. (1978a). "Stimulus dominance and ear dominance in the perception of dichotic voicing contrasts," Brain Lang. 5, 310–330.

Repp, B. H. (1978b). "Stimulus dominance in fused dichotic syllables," Haskins SR-55/56, 133–148.

Repp, B. H. (1978c). "Stimulus dominance and ear dominance in fused dichotic speech and nonspeech stimuli," Haskins SR-55/56, 163–180.

Repp, B. H. (1979). "Relative amplitude of aspiration noise as a voicing cue for syllable-initial stop consonants," Lang. Speech 22, 173–189.

Speaks, C., and Bissonette, L. J. (1975). "Interaural-intensive differences and dichotic listening," J. Acoust. Soc. Am. 58, 893–898.

Studdert-Kennedy, M., Shankweiler, D., and Pisoni, D. B. (1972). "Auditory and phonetic processes in speech perception: Evidence from a dichotic study," Cogn. Psychol. 3, 455–466.

Whittaker, R. G., and Porter, R. J., Jr. (1976). "Dichotic and monotic masking of CV's by second formants with different transition starting values," J. Acoust. Soc. Am. Suppl. 1 60, S119(A).