STIMULUS DOMINANCE AND EAR DOMINANCE IN FUSED DICHOTIC SPEECH AND NONSPEECH STIMULI: A REPLICATION

Bruno H. Repp

Abstract. The patterns of stimulus dominance and ear dominance effects were compared between two types of fused dichotic stimuli: two-formant CV syllables ranging from /bV/ to /dV/ to /gV/ (V = /a/ or /i/) and brief, isolated, steady-state resonances ("timbres") corresponding to the second-formant onset frequencies of the CV stimuli. Results for the two types of stimuli were similar; there was no evidence that the speechlike quality of the CV syllables influenced stimulus dominance or ear dominance effects, which thus seemed to be governed by auditory stimulus properties and individual differences in their perception. This result confirms earlier data of Repp (1978c).

INTRODUCTION

The present study replicates and extends an earlier experiment (Repp, 1978c), and the reader is referred to that report for a general introduction. Briefly, the purpose of the earlier study was to investigate whether the relative speechlikeness of a set of stimuli influences the pattern of stimulus dominance and ear dominance effects obtained in fused dichotic presentation. The answer was negative—stimulus dominance seemed to be governed by auditory stimulus properties (second-formant onset frequency), and the direction of ear dominance depended primarily on the individual listener, not on stimulus type. These findings provided evidence against the hypotheses that dichotic stimulus dominance reflects the relative "category goodness" of speech stimuli (Repp, 1976, 1977a, 1978a, 1978b), and that, within the range of stimuli considered, the right ear would become more dominant as the stimuli become more speechlike.

There were five types of synthetic stimuli in the earlier study, all derived from two-formant syllables ranging from /bæ/ to /dæ/ to /gæ/. Due to the small number of subjects and the difficulty of the task, the results for three of the stimulus types ("bleats", "transitions", and "chirps") were preliminary at best. However, more stable data were available for the two extreme (i.e., respectively most and least speechlike) stimulus sets—full CV syllables and "timbres" (40-msec steady-state, second-formant resonances at the frequencies that, as onset values, distinguished the CV syllables). It was the comparison between these, acoustically most dissimilar, sets that revealed the greatest similarity in response patterns.

Acknowledgment: This research was supported by NICHD Grant HD01994 and BRS Grant RR05596.

[HASKINS LABORATORIES: Status Report on Speech Research SR-59/60 (1979)]
CV syllables and timbres had in common the distinctive property of second-formant (F₂) onset frequency. The similarity of results for the two types of stimuli suggested that both stimulus dominance and ear dominance effects were governed by this auditory stimulus property (as well as by individual differences in its perception). To test the generality of this finding, the present study attempted to replicate it with two new sets of CV syllables and timbres. The new stimuli represented different ranges of F₂ onset frequencies, and, in the CV syllables, different vocalic contexts (/a/-/i/) than in the earlier stimuli (/a/-/æ/). Shifts up or down the frequency scale were not expected to affect responses to steady-state timbres; this part of the study was almost certainly expected to replicate the earlier results. In CV syllables, on the other hand, the F₂ onset frequency changes rapidly toward a steady state characteristic of the vowel; depending on this vowel, a given onset frequency may initiate a rising, flat, or falling formant transition. If all that mattered in dichotic competition were F₂ onset frequency, then, of course, the nature of the transition and of the following vowel would be irrelevant. However, Pomponio, Rilhac-Sutter, Simon, and Sommer (1977)—in a study that almost duplicates the CV condition of the present experiment but only recently came to my attention—report that the steepness of F₂ transitions is the decisive factor in dichotic competition. If so, quite different patterns of stimulus dominance relationships would be expected for the two sets of CV syllables described below. Moreover, it is well known that the perceived place of articulation of syllable-initial stop consonants is not determined by formant onset frequency alone, and to the extent that phonetic categorization influences dichotic stimulus dominance (as postulated by the "category goodness hypothesis" of Repp, 1976, 1977a, 1978a, 1978b), the two sets of CV syllables to be described should yield different response patterns. On the other hand, identical response patterns for all four stimulus sets (two CV series and two timbre series) would provide strong support for the sole importance of F₂ onset frequency in dichotic competition.

The predictions just discussed concern the relative dominance of one stimulus over another in dichotic competition. As to ear dominance, the questions were: Does ear dominance shift toward the right side (left hemisphere) as the stimuli become more speechlike? And are ear dominance effects for CV syllables and timbres related to each other? Repp's (1978c) earlier data suggested a negative answer to the first question and a positive answer to the second one, but because of the small number of subjects, a replication seemed desirable.

Method

Subjects. The eight subjects included six paid student volunteers, an undergraduate research assistant, and the author. All subjects had listened to synthetic speech before, but only the author had had extensive experience. A full replication of the author's data was available; these two sets of data were averaged before they were combined with those of the other subjects.

Stimuli. There were four sets of stimuli: /Ca/ syllables, /Ci/ syllables, and two corresponding sets of timbres. All stimuli were generated on the Haskins Laboratories parallel resonance synthesizer. Each set contained seven stimuli distinguished by F₂ onset frequency. The CV syllables were wholly periodic two-formant patterns with initial stepwise-linear formant
transitions that led to the perception of /b,d,g/ preceding either /a/ or /i/. The \( F_1 \) transition was constant and 30 msec in duration; it went from 181 Hz to 743 Hz in /Ca/ syllables and from 181 Hz to 286 Hz in /Ci/ syllables. The \( F_2 \) transitions were 40 msec in duration and ended at 1075 Hz in /Ca/ stimuli and at 2307 Hz in /Ci/ stimuli. The variable \( F_2 \) onset frequencies are shown in Table 1. The portion following the formant transitions was perfectly steady-state, and total stimulus duration was 250 msec. To increase their speechlike quality, the steady-state portions of the CV syllables were given a linearly falling \( F_0 \) contour (from 114 Hz to 90 Hz), whereas the initial transitional portions had a constant \( F_0 \) of 114 Hz.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>/Ca/ syllables</th>
<th>/Ci/ syllables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low timbres</td>
<td>High timbres</td>
</tr>
<tr>
<td>1</td>
<td>921</td>
<td>1541</td>
</tr>
<tr>
<td>2</td>
<td>1075</td>
<td>1695</td>
</tr>
<tr>
<td>3</td>
<td>1232</td>
<td>1845</td>
</tr>
<tr>
<td>4</td>
<td>1386</td>
<td>1996</td>
</tr>
<tr>
<td>5</td>
<td>1541</td>
<td>2156</td>
</tr>
<tr>
<td>6</td>
<td>1695</td>
<td>2307</td>
</tr>
<tr>
<td>7</td>
<td>1845</td>
<td>2462</td>
</tr>
</tbody>
</table>

The two sets of timbres were matched to the two sets of CV stimuli. Timbres consisted of 40-msec steady-state \( F_2 \) segments at the frequencies listed in Table 1, with a constant \( F_0 \) of 114 Hz. Since the \( F_2 \) onset frequencies were lower in /Ca/ syllables than in /Ci/ syllables, the corresponding sets of timbres will be referred to as "low" and "high," respectively. Note that stimuli 5-7 of the low set were identical with stimuli 1-3 of the high set.

All stimuli were digitized at 10 kHz using the Haskins Laboratories pulse code modulation system. There were four experimental tapes, one for each of the four stimulus sets. Each tape contained first a binaural AXB discrimination test in which the five two-step pairings (1-3, 2-4, 3-5, 4-6, 5-7) of the stimuli within a set occurred three times in each of the four AXB configurations (AAB, ABB, BAA, BBA), yielding a total of 60 triads. This sequence was followed by a dichotic AXB test. In each dichotic AXB triad, X consisted of A and B presented simultaneously to the two ears, which resulted in strong fusion; A and B were presented binaurally. All fifteen pairings of stimuli two or more steps apart in a given series appeared in each of the four possible AXB configurations (A[AB]B, A[B]BA, B[AB]A, B[BA]A), yielding a total of 60 triads. Three different randomizations of these 60 triads were recorded. Finally, for each of the two CV syllable sets, there was a standard identification test containing twenty replications of each of the seven stimuli. The interstimulus intervals in this test, and between AXB triads,
were 3 sec; those within AXB triads were 500 msec.

Procedure. Each subject listened to each of the four tapes in four separate sessions. The order was counterbalanced across subjects. The order of tests within conditions was as described above. The main purpose of the binaural AXB discrimination test was to familiarize the subjects with the stimuli. The task was to write down "A" when X equaled A, and "B" when X equaled B. In the dichotic AXB test, the task was the same, except that "was more similar to" replaced "equalled" in the instructions. The three dichotic sequences were repeated after a pause during which the channels-to-ears assignment was electronically reversed. Thus, each subject gave a total of 24 responses to each dichotic stimulus combination (disregarding ear assignment). In the final identification test (CV syllables only), the subjects identified the initial consonant in each stimulus as B, D, or G. Other details of procedure were the same as in Repp (1978c), as indeed were most of those described above.

Results and Discussion

Binaural identification and discrimination. The labeling and two-step discrimination results for the stimuli in the two speechlike series, averaged over subjects, are shown in Figure 1. It can be seen that the endpoint stimuli of the CV continua were rather consistently labeled B and G, respectively. However, no stimulus on either continuum was consistently labeled D, although D responses reached a maximum for stimuli 4 and 5. Presumably, alveolar stops need a third-formant transition or a burst (both absent in the present stimuli) to sound convincing. Two-step discrimination performance ranged from 62 to 90 percent. The precise pattern of discrimination results need not concern us here, but note that the discriminability of the individual stimuli set an upper limit to performance in the dichotic AXB test. However, that test contained not only two-step stimulus combinations—which may have presented some difficulty for the listeners—but also combinations of stimuli three to six steps apart; these, of course, were much easier to discriminate.

No labeling data were collected for the timbres, since they do not fall into natural categories. AXB discrimination performance for timbres was virtually perfect; only one subject made any errors at all.

Dichotic stimulus dominance. The average results of the dichotic AXB tests are plotted in Figure 2 as "percent 'i-ness' judgments (i<j)", i.e., the percentage of trials on which a fused dichotic stimulus was judged to be more similar to the component with the lower $F_2$ onset frequency. Individual stimulus pairs are identified by stimulus numbers in the graphs (i) and on the abscissa (j). Let us focus first on the results for timbres, shown on the left. First, it is evident that nearly all data points fall above the 50-percent (equilibrium) line. This indicates a strong tendency for low-frequency timbres to dominate high-frequency timbres, which replicates a similar trend found by Repp (1978c). As in the earlier study, this trend was shown by most, but not all, subjects. One subject, in particular, showed exactly the opposite (viz., high-frequency dominance), and to such an extent that her timbre data could not be used. Moreover, a second subject showed such strong low-frequency dominance for the set of low timbres that his data provided no information and were likewise excluded from the averages. Thus,
Figure 1. Labeling and 2-step discrimination functions for /Ca/ and /Ci/ stimuli.
the low timbre data in Figure 2 are based on only six subjects, whereas the high timbre data derive from seven subjects.

A second feature of the data to note is that stimulus pair 5-7 in the low set and stimulus pair 1-3 in the high set were physically identical but showed quite different results. In general, the fact that approximately equal degrees of overall low-frequency dominance were obtained in both sets of timbres (and in the timbre stimuli of Repp, 1978c) demonstrates that the degree of frequency dominance for a given stimulus pair was highly range-dependent. Obviously, listeners adapted to the frequency range of a given stimulus set and adjusted their criteria accordingly.

Third, we note that the main determinant of stimulus dominance within a given stimulus set was the frequency of the lower timbre in a pair: the lower its frequency, the more dominant it became. This is indicated by the vertical separation of the functions in Figure 2, which connect all points with a constant lower timbre. On the other hand, the frequency of the higher timbre in a pair had a much smaller effect, as indicated by the shallow slopes of the functions in Figure 2. Whatever effect there was, it too was in the direction of increased dominance as frequency changed from high to low.

Because of their peculiar structure, these data are difficult to analyze statistically. However, the consistency of the average response pattern is expressed in the (within-group) correlation across the two sets of timbres $r = +0.84, p < .001$, and in their respective (between-group) correlations with the timbre data of Repp (1978c), obtained with stimuli in a frequency region precisely intermediate between the present two ranges ($r = +0.83$ and $+0.88$, both $p < .001$).

Turning now to the results for CV syllables, shown in the right half of Figure 2, we find a rather similar pattern. Overall low-frequency dominance was less pronounced, and so was the effect of the frequency of the stimulus with the lower $F_2$ onset, but both trends were clearly present, especially in /CI/ syllables. The frequency of the higher-onset stimulus seemed to have no effect at all here (or, in /CA/ stimuli, perhaps in a direction opposite to that observed in timbres). In general, the results for /CA/ syllables showed no strong effects of any sort, indicating difficulties in discriminating and judging these stimuli. One factor that may have contributed to these difficulties was the tendency of simultaneous /ba/ and /ga/ to yield /da/ (cf. Cutting, 1976; Repp, 1976); this tendency was much less pronounced in /CI/ stimuli, according to the author's observations as a subject. Of course, this phenomenon introduced random responses, since listeners did not (and in fact, had been encouraged not to) consistently judge /da/ to be more similar to either /ba/ or /ga/.

The important point about the CV data is that they provide little indication of a response pattern specific to their speechlike quality; in this respect, they replicate the data of Repp (1978c). From the category goodness hypothesis, one should have expected stimulus 7 (a good exemplar of /gi/ and a reasonable one of /ga/) to be a much stronger competitor than stimuli 3-6. There was a small trend in that direction—too small to be taken seriously. Nor do the data confirm the observation of Pompino et al. (1977) that the steepness of transitions determines their relative dominance. The reason for this discrepancy is not quite clear. However, the present results do support,
Figure 2. Stimulus dominance patterns for four types of stimuli.
in a rather indirect way, the suggestion (e.g., Stevens & Blumstein, 1978) that listeners are most sensitive to the onset spectrum of CV syllables.

The relative consistency of the CV data is indicated by the correlation (across the 15 data points) between the /Ca/ and /Ci/ results ($r = +0.73$, $p < .001$), as well as by at least one of their respective correlations with the earlier /Ca/ data ($r = +0.36$, n.s., and $r = +0.78$, $p < .001$). Moreover, the response patterns for CV syllables and timbres were significantly similar in the present study (correlations between +0.48 and +0.82), as they had also been in Repp (1978c). Thus, it may be concluded that the stimulus dominance patterns for all stimulus series were essentially the same, all surface differences probably being due to variations in task difficulty.

**Ear dominance effects.** Individual ear dominance effects for the four stimulus sets are displayed in Table 2. The index shown is $e'$ (Repp, 1977b), which ranges from -1 for perfect left-ear dominance to +1 for perfect right-ear dominance. As can be seen, there was a considerable number of significant effects, both in favor of the left ear and in favor of the right ear. What is not evident in the data is a consistent change in the direction of ear dominance as a function of stimulus type. Moreover, the ear dominance effects for all sets of stimuli seemed to be related (correlation coefficients between +0.51 and +0.78). Thus, the relative speechlikeness of the stimuli appeared to have no systematic influence on ear dominance. This once more confirms Repp (1978c).

---

**Table 2**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Low timbres</th>
<th>High timbres</th>
<th>/Ca/ syllables</th>
<th>/Ci/ syllables</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a (I)</td>
<td>+0.02</td>
<td>+0.81***</td>
<td>+0.34***</td>
<td>+0.26**</td>
</tr>
<tr>
<td>(II)</td>
<td>+0.20**</td>
<td>+0.56***</td>
<td>+0.16</td>
<td>+0.58***</td>
</tr>
<tr>
<td>2</td>
<td>+0.12**</td>
<td>+0.29**</td>
<td>-0.04</td>
<td>-0.41***</td>
</tr>
<tr>
<td>3</td>
<td>+0.21</td>
<td>-0.33***</td>
<td>+0.01</td>
<td>+0.06</td>
</tr>
<tr>
<td>4</td>
<td>+0.01</td>
<td>-0.14</td>
<td>-0.08</td>
<td>+0.11</td>
</tr>
<tr>
<td>5</td>
<td>-0.38**</td>
<td>-0.14</td>
<td>-0.11</td>
<td>-0.13*</td>
</tr>
<tr>
<td>6</td>
<td>-0.46***</td>
<td>-0.71***</td>
<td>-0.14*</td>
<td>-0.08</td>
</tr>
<tr>
<td>7</td>
<td>----b</td>
<td>----b</td>
<td>+0.14*</td>
<td>-0.05</td>
</tr>
<tr>
<td>8</td>
<td>-0.41***</td>
<td>-0.77***</td>
<td>-0.27**</td>
<td>-0.55***</td>
</tr>
</tbody>
</table>

***$p < .001$
**$p < .01$
* $p < .05$

a The author; data from two sessions.
b No estimate because of extreme stimulus dominance effects.

---

The author, who was the only subject to show consistent right-ear dominance in all four sets of stimuli, also happened to be the only strongly
right-handed subject. By a curious coincidence, all other subjects (drawn from a limited population of summer students) were either left-handed or had left-handed relatives. Perhaps, then, the hypothesis that ear dominance would be shifted to the right in CV syllables, relative to timbres, was not given a proper test, as some of the subjects may not have been left-hemisphere-dominant for speech. The subjects most likely to fall in that category were subjects 6 and 8, who were left-handers with left-handed relatives (Hardyck & Petrinovich, 1977). These subjects did show the largest left–ear dominance effects in the group; however, contrary to expectations, they were also the ones who showed most clearly a reduction of left–ear dominance with CV syllables—the opposite of what one should have expected if these subjects were right–hemisphere-dominant for speech. Of course, a reduction in absolute ear dominance was to be expected for CV syllables because of the greater difficulty of that condition. But subjects were not even consistent in that respect. Thus, the data continue to offer no consistent evidence of systematic variation in ear dominance among stimulus conditions. They are perhaps suggestive of a relation between ear dominance (for both timbres and CV syllables) and hemispheric dominance for speech, but this issue clearly requires further investigation, as does the possible relation between ear dominance for timbres and for pure tones contrasting in pitch. (Efron & Yund, 1974).

Conclusion

The present study confirms that of Repp (1978c) in all aspects. It provides no convincing evidence that the relative speechlikeness of fused complex sounds contrasting in harmonic spectrum influences either stimulus dominance or ear dominance effects. Stimulus dominance seems to be a function of F₂ onset frequency and of the frequency range employed. Ear dominance depends primarily on the individual listener and on certain task variables, not on stimulus type; its relation to cerebral dominance for speech and/or pitch perception remains to be established.

REFERENCES


