Perception & Psychophysics
1990, 48 (2), 169-178

Interaction among auditory dimensions: Timbre, pitch, and loudness

ROBERT D. MELARA and LAWRENCE E. MARKS
John B. Pierce Foundation Laboratory and Yale University, New Haven, Connecticut

In two experiments, we examined whether or not pairs of auditory dimensions—timbre-loudness (Experiment 1) and timbre-pitch (Experiment 2)—interact in speeded classification. Subjects classified values from one dimension while the other dimension was (1) held constant (baseline), (2) varied orthogonally (filtering), or (3) correlated linearly. The subjects showed substantial Garner interference when classifying all dimensions—that is, poor performance at filtering relative to baseline. Timbre and loudness displayed redundancy gain (i.e., performance faster than baseline) when correlated positively, but redundancy loss (i.e., interference) when correlated negatively. Timbre and pitch displayed redundancy gain however dimensions were correlated. Both pairs of dimensions showed substantial effects of congruity: Attributes from one dimension were classified faster when paired with "congruent" attributes from the other dimension. The results are interpreted in terms of an interactive multichannel model of auditory processing.

How do perceptual systems analyze input from separate dimensions of sensory stimulation? The present study represents a continuation of our investigations into the mechanisms of multidimensional processing (Melara, 1989a, 1989b; Melara & Marks, 1990a, 1990b, 1990c, in press; Melara & O'Brien, 1987). Guiding our research is a general model of how any set of multidimensional stimuli is processed, regardless of whether stimuli trigger visual, auditory, or other sensory receptors. Specifically, we conceive of stimulus processing in terms of a bank of higher order filters or channels. As we discuss elsewhere (see Melara & Marks, 1990c), each channel is associated with a psychologically meaningful dimension that we call the *primary* dimension.[1] Such primary perceptual dimensions vary maximally with changes in particular physical dimensions (e.g., frequency, intensity, or duty cycle). Thus, in audition, separate channels code changes in pitch (frequency), loudness (intensity), and timbre (duty cycle), in addition to changes occurring along other dimensions such as duration.

Because of perceptual primacy, the perceiver is able, through attentional control, to select and then access the results of analyses in a given channel (e.g., for purposes of classification or identification). We suggest that channel identification (e.g., selecting the pitch channel to make frequency discriminations) is reasonably accurate for all primary dimensions, independent of changes occurring in other channels. The important factor in multidimensional processing, we hypothesize, is how the output from one channel (e.g., pitch) is weighted by output from other channels (e.g., loudness).

A given pair of dimensions may be either *separable* or *interacting* (see Garner, 1974, 1981; Lockhead, 1966, 1972, 1979). If two dimensions are separable, subjects are able to attend selectively to either dimension, ignoring orthogonal variation on the irrelevant dimension. Circle size and diameter orientation are an example of separable dimensions (Garner & Felfoldy, 1970). In our view, with separable dimensions, the output received from the selected channel (dimension) is perceptually independent (cf. Ashby & Townsend, 1986) of the output from the irrelevant channel. This situation is depicted graphically for the dimensions of size and orientation in panel A of Figure 1.

If two dimensions interact, orthogonal variation on the irrelevant dimension causes *Garner interference* (see Pomerantz, 1983, 1986; Pomerantz, Pristach, & Carson, 1989) when the relevant dimension is being classified. Interaction is therefore indicated by a failure of selective attention (see, e.g., Garner, 1974, 1981). We stress, however, that although perceivers fail to attend selectively to interacting dimensions, they can nonetheless select the appropriate channel for classification (i.e., channel identification is accurate). Saturation and brightness of colors are an example of interacting dimensions (Garner & Felfoldy, 1970).

We hypothesize that, with interacting dimensions, channels are linked at some level (or levels) of information processing, be that level sensory/perceptual, phonemic/graphemic, lexical, or semantic. Interaction is therefore defined by us as the leakage or *crosstalk* between channels at a particular level of processing of each dimension.

Preparation of this article was supported by National Institutes of Health (NIH) Postdoctoral Fellowship HD07137 and NIH Grant NS28617 to Robert Melara and by NIH Grant DC00271 and National Science Foundation Grant BNS84-20017 to Lawrence Marks. We thank Edward Carterette, Bonnie Potts, and two anonymous reviewers for their comments on an earlier version. We also thank Ed Washecka, Trex Proffitt, and Alison Wittenberg for their assistance in collecting and analyzing the data. Address all correspondence to Robert D. Melara, who is now in the Department of Psychological Sciences at Purdue University, West Lafayette, IN 47907.

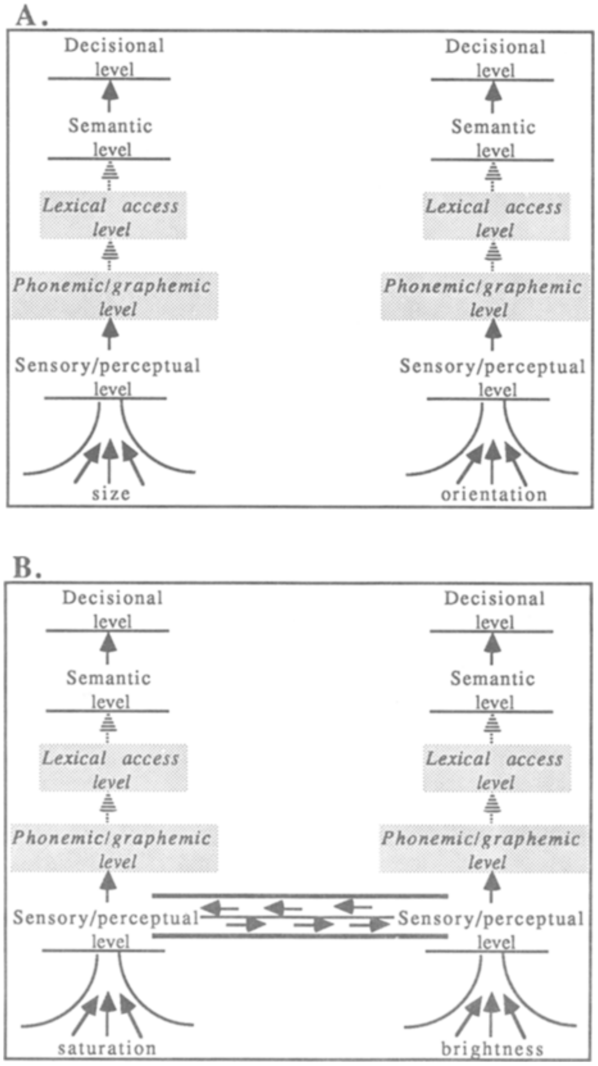

Figure 1. Pictorial interpretation of processes underlying the separability of circle size and diameter orientation (A) and those underlying the integrality of saturation and brightness (B). Arrows indicate the hypothesized flow of processing from one level of analysis to another, either during analyses of any one dimension or when crosstalk occurs between two dimensions. Labels have been masked for the linguistic analyses (i.e., lexical access level, phonemic/graphemic level), which are presumed to be disengaged when processing these perceptual dimensions.

According to our model, the partial output from the irrelevant (unselected) channel influences what perceptual values (attributes) are derived as final output of processing in the selected channel (Melara & Marks, 1990b). Garner interference may be either symmetric (present when either dimension is being classified) or asymmetric (present for only a single dimension), depending on what two dimensions are paired; theoretically, then, crosstalk may be either bidirectional (mutual) or unidirectional (see Melara & Marks, 1990a). Crosstalk may occur either within a single level of processing or between two different levels. An example of bidirectional crosstalk within the sensory/perceptual level is depicted for saturation and brightness in panel B of Figure 1.

Are dimensions of sound separable or interacting? Research on this question has been confined primarily to the auditory dimensions of pitch and loudness. Evidence indicates that pitch and loudness do interact: In speeded classification, subjects' performance suffers substantial Garner interference when pitch and loudness are varied orthogonally (Grau & Kemler-Nelson, 1988; Melara & Marks, 1990c; Wood, 1975). Moreover, performance benefits when these two dimensions are correlated (e.g., high pitch paired with loud and low pitch paired with soft); that is, classification speed and accuracy show a *redundancy gain* (Grau & Kemler-Nelson, 1988; Melara & Marks, 1990c). These results are consistent with the view that crosstalk occurs between dimensional channels at some level, presumably within sensory/perceptual processing (see Grau & Kemler-Nelson, 1988; Melara & Marks, 1990c; Wood, 1975; see Figure 2).

Relatively little research has been devoted toward understanding the information processing characteristics of auditory timbre (but see Iverson & Krumhansl, 1989). Timbre is the tonal quality of complex sounds. It is the characteristic of sound that permits us to distinguish a guitar note from a piano note played at the same fundamental frequency and intensity. The lack of research aimed at evaluating timbre classification reflects, perhaps, the inherent difficulty in adequately defining values along this dimension (see below). Nonetheless, a gap presently exists in our understanding of multidimensional processing with the three primary dimensions of sound—timbre, pitch, and loudness. The present study represents our initial attempt to close this gap by examining timbre classification when that dimension is paired with either pitch or loudness.

How perceivers process timbre—specifically, whether timbre interacts with or is separable from dimensions of pitch and loudness—bears heavily on the development of models of auditory processing. A number of possibilities exist. One is that timbre interacts bidirectionally with both dimensions. Here, one might envision a triplex of parallel channels, interconnected at one level (or perhaps more); later, we shall discuss just such a model. Another possibility is that timbre is separable from one or both dimensions. Evidence for separability could suggest that the auditory system carries out analyses sequentially on separate dimensions, as might occur if profile analyses (cf. Green, 1988) followed encoding of frequency and amplitude. A third possibility is that crosstalk is unidirectional; such a result could also be interpreted to reflect dimensional hierarchies in auditory processing. Discrimination among these possibilities—that is, determination of the relationships among timbre, pitch, and loudness—has obvious

implications for the development of psychological and psychophysical models of broadband signal analysis (e.g., Durlach, Braida, & Ito, 1986).

A final reason for examining auditory interactions is to pursue further our distinction between *single context* and *dual context* (Melara & Marks, 1990b), and the differential effects these two contexts have on classification. Single context occurs when the irrelevant dimension is varied under a single set of constraints. For example, varying two dimensions orthogonally, as in "filtering" tasks (Garner, 1974; Posner, 1964), produces an "intraclass context," because stimuli vary within each response category. Dual context occurs when two forms of variation are combined. Thus, if stimuli vary within each category (intraclass context) but, in addition, the relevant and irrelevant dimensions are correlated ("redundant context"), a dual context is created.

We have found interesting differences between single contexts and dual contexts in the classification of auditory dimensions. When pitch and loudness are paired, for example, judgments on both dimensions show interference with single (intraclass) contextual variation. With dual (intraclass and redundant) contextual variation, however, only loudness classifications are worse than baseline; pitch judgments do not suffer from intraclass context, but show benefit from redundant context. On the basis of this evidence, we call pitch a HARD dimension, because in dual context intraclass interference can be avoided, and we call loudness a SOFT dimension, because in dual context intraclass interference is inevitable.

Results from dual context also identify timbre as a HARD dimension: Classification of this dimension shows no interference when the dimension is paired with either pitch or loudness (Melara & Marks, 1990b). But note that the absence of interference can also obtain if timbre is separable from pitch and loudness. Thus, it is critical to evaluate whether timbre interacts with these dimensions when single context is established—that is, whether timbre judgments show Garner interference in speeded classification. That was our purpose in carrying out the present experiments.

In our investigation of dual context, we defined timbre as the duty cycle of a variable pulse (rectangular) tone. Timbre can be defined in terms of either spectral properties—that is, the distribution of harmonics of the fundamental frequency—or dynamic properties—that is, the change in harmonic structure across time. In the present study, as in the dual context study, we manipulated only spectral timbre (i.e., duty cycle). We were therefore able to make direct comparisons between the effects on spectral timbre of single and dual context when context is created by auditory pitch and loudness.

We paired timbre with loudness (intensity) in Experiment 1, and timbre with pitch (frequency) in Experiment 2. In each experiment, we employed the full speeded classification procedure developed by Garner (1974; Garner & Felfoldy, 1970). In this procedure, one dimension is criterial (at any one time), and the other dimension is irrelevant. Subjects identify one of two possible values along the criterial dimension. Meanwhile, the irrelevant dimension is manipulated in one of three ways: (1) It is held at a single value in *baseline* tasks; (2) it is varied in an uncorrelated fashion with the criterial dimension in *filtering* tasks; (3) it is correlated perfectly with the criterial dimension in *correlated* tasks. Garner interference occurs if performance is inferior in filtering tasks relative to baseline. Redundancy gain occurs if performance is superior in the correlated tasks relative to baseline. In each experiment, we ask: Do pairs of auditory dimensions yield Garner interference and/or redundancy gains in speeded classification?

In the present study, we also tested for effects of dimensional correspondence, a phenomenon we have examined in several intramodal and intermodal pairs of dimensions (e.g., Marks, 1987; Melara, 1989a; Melara & Marks, 1990a, 1990c, in press; Melara & O'Brien, 1987; see also Clark & Brownell, 1975, 1976; Pomerantz, 1983; Walker & Smith, 1984, 1986). In particular, previous research with pitch and loudness (Melara & Marks, 1990c) revealed substantial *congruity effects*: Subjects were faster to respond to high-pitched loud sounds or low-pitched soft sounds than to high-pitched soft sounds or low-pitched loud sounds. It is possible that pitch-loudness congruity arises because of a *polar* correspondence of attributes: The congruent stimuli comprise attributes from either the dimensions' positive poles (i.e., high pitch and loud) or their negative poles (i.e., low pitch and soft). According to a polar account, then, perceivers link the corresponding poles of interacting dimensions, perhaps at a sensory/perceptual level, and congruity effects arise as subjects work to overcome mismatching poles, perhaps at a decision or response stage. The present study permitted us to explore this idea further.

In these experiments, we began with an arbitrary definition of *congruent-incongruent*. Sounds having a relatively long duty cycle ("hollow" sounds) were assigned to one pole, and sounds having a short duty cycle ("twangy" sounds) were assigned to the other pole. Then, in Experiment 1, we arbitrarily designated hollow-soft and twangy-loud sounds as congruent, and hollow-loud and twangy-soft sounds as incongruent. For each subject, we derived congruity scores—that is, the arithmetical reaction time (RT) difference between sounds having congruent attributes and those having incongruent attributes. If, during classification, pairs of sound attributes are processed for their polar correspondence or noncorrespondence, then a significant congruity effect should obtain: Sounds with congruent attributes should be classified faster than sounds with incongruent attributes, leading to an overall congruity score significantly different from zero. If the congruity effect is significantly positive, it implies that our designation of congruent-incongruent matched that of our subjects. If the congruity effect is significantly negative, it implies that our coding was opposite to that of our sub-

jects. As we shall see, the pattern of congruity effects obtained here complicates a simple polar correspondence model of auditory congruity effects.

EXPERIMENT 1
TIMBRE AND LOUDNESS

Method

Subjects. Twenty subjects (8 men and 12 women), mostly students recruited from the Yale community, were paid $5 per hour to participate.

Stimuli and Apparatus. Signals were generated by the SID chip of a Commodore 64 microcomputer, directed to a Realistic SA-150 stereo amplifier, and presented over Realistic Nova-40 headphones. The stimulus set contained four sounds—namely, two timbres at each of two intensities. The two values on the timbre (duty cycle) dimension were .1878 (twangy) and .3128 (hollow), where .50 represents a square wave. We combined each of these with two sound intensities, 64 dB (soft) and 70 dB (loud), calibrated on the A-scale of a General Radio 1565-B sound level meter. All four sounds (i.e., hollow-soft, hollow-loud, twangy-soft, twangy-loud) were complex rectangular waveforms (rise and decay times = 10 msec) having identical fundamental frequencies of 950 Hz. Power spectra for the two signal waveforms delivered to the headphones appear in panel A (.1878 duty cycle) and panel B (.3128 duty cycle) of Figure 2. Power is represented up to and including the 15th harmonic, which, in each case, accounted for over 99% of total power in the sound.

From the four stimuli, we created 10 experimental tasks, 5 involving timbre classification and 5 involving loudness classification. Each task contained 48 trials. Four tasks were baseline, 2 involved positively correlated dimensions, 2 involved negatively correlated dimensions, and 2 involved orthogonally varied dimensions (filtering).

Baseline. Subjects discriminated between the two values on one dimension (e.g., hollow vs. twangy timbre), while the other dimension was held constant at one of its two possible values (e.g., soft).

Positively correlated dimensions. Subjects discriminated the two congruent stimuli from each other—that is, the stimuli whose values came from corresponding poles of their respective dimensions (i.e., hollow-soft and twangy-loud). Subjects discriminated values along either the timbre or the loudness dimension.

Negatively correlated dimensions. Subjects discriminated the two incongruent stimuli (i.e., hollow-loud and twangy-soft) from each other, according to either timbre or loudness.

Filtering. Subjects classified all four stimuli into either two timbre categories (hollow vs. twangy) or two loudness categories (loud vs. soft).

Equal numbers of the two stimuli (baseline, correlated dimensions) or four stimuli (filtering) were included in each task.

Procedure. The subjects were tested individually in a darkened sound-attenuating booth. The subjects performed the five timbre or loudness tasks together as a set, with half of the subjects performing timbre judgments first and half performing loudness judgments first. Order of task within each set of five followed four 5 × 5 Latin squares. Each of the two sets of judgments began with 48 trials of practice in a baseline task. No practice data were analyzed.

Trials were presented randomly. Each trial began when the subject pressed the space bar on the computer's keyboard. The stimulus appeared 1 sec later. Discriminations were made by pressing either a right-hand or a left-hand key on the keyboard. Key assignment was counterbalanced across subjects. The elapsed time between stimulus onset and keypress was measured in milliseconds, using a software timer (Price, 1979). The subjects were instructed to make their decisions as quickly as possible without error. A message displayed on the computer's monitor informed the subjects of incorrect responses. Subjects responding slower than 800 msec were encouraged (through a screen message) to respond more quickly. Following each task, the subjects were informed of their mean RT and error rate. The entire experiment lasted approximately 45 min.

Results

Overall performance. Mean RTs and error rates for each task appear in Table 1. The subjects discriminated loudnesses (346 msec) 29 msec faster on the average than they discriminated timbres [375 msec; $F(1,19) = 8.47$, $p < .01$], indicating that baseline discriminabilities were uneven and in favor of loudness. The overall error rate was .06. RTs correlated .87 with errors.

Effects of correlated and uncorrelated dimensions. This experiment resulted in 51 msec of Garner interference; that is, subjects were 51 msec slower at filtering (irrelevant dimension varying) than baseline [irrelevant dimension constant; $F(1,19) = 41.48$, $p < .001$]. The effect was significant for each dimension tested separately [timbre judgments, $F(1,19) = 18.39$, $p < .001$; loudness judgments, $F(1,19) = 58.97$, $p < .001$], and there was no interaction between dimension (timbre, loudness) and task [baseline, filtering; $F(1,19) = .89$, n.s.], indicating that the effect was equivalent for both dimensions.

Positively correlated dimensions were classified 36 msec faster than baseline [$F(1,19) = 30.41$, $p < .001$], which was a redundancy gain, but negatively correlated dimensions were classified 24 msec slower than baseline [$F(1,19) = 11.5$, $p < .01$], which was a redundancy *loss* (see Pomerantz, 1986). Again, these results were consistent across dimensions: timbre—gain from positive correlation [$F(1,19) = 36.47$, $p < .001$]; loss from negative correlation [$F(1,19) = 11.88$, $p < .01$]; loudness—gain from positive correlation [$F(1,19) = 52.35$, $p < .001$]; loss from negative correlation [$F(1,19) = 7.26$, $p = .01$]. There were no significant interactions.

Effects of congruity. An ANOVA revealed that stimuli containing congruent attributes were classified 30 msec faster than stimuli containing incongruent attributes [$F(1,19) = 79.27$, $p < .001$]. Congruity scores—that is, differences between mean RTs to incongruent and congruent stimuli, were derived for each subject for baseline, correlated dimensions, and filtering classifications. These appear in Table 2. Individual t tests were used to indicate whether these scores were reliably greater than zero (i.e., a congruity effect). The congruity effects were significant in each of the three types of classification: baseline [6 msec; $t(19) = 2.28$, $p = .03$]; correlated dimensions [60 msec; $t(19) = 8.85$, $p < .001$]; filtering [18 msec; $t(19) = 3.61$, $p = .002$]. An ANOVA and a subsequent Newman-Keuls post hoc test indicated that the congruity effect for correlated dimensions was significantly larger than that for either baseline or filtering, which did not differ from each other [$F(2,38) = 36.33$, $p < .001$]. In addition, we found a significant interaction between dimension (timbre, loudness) and task [baseline, filtering, correlated dimensions; $F(2,38) = 7.72$, $p =$

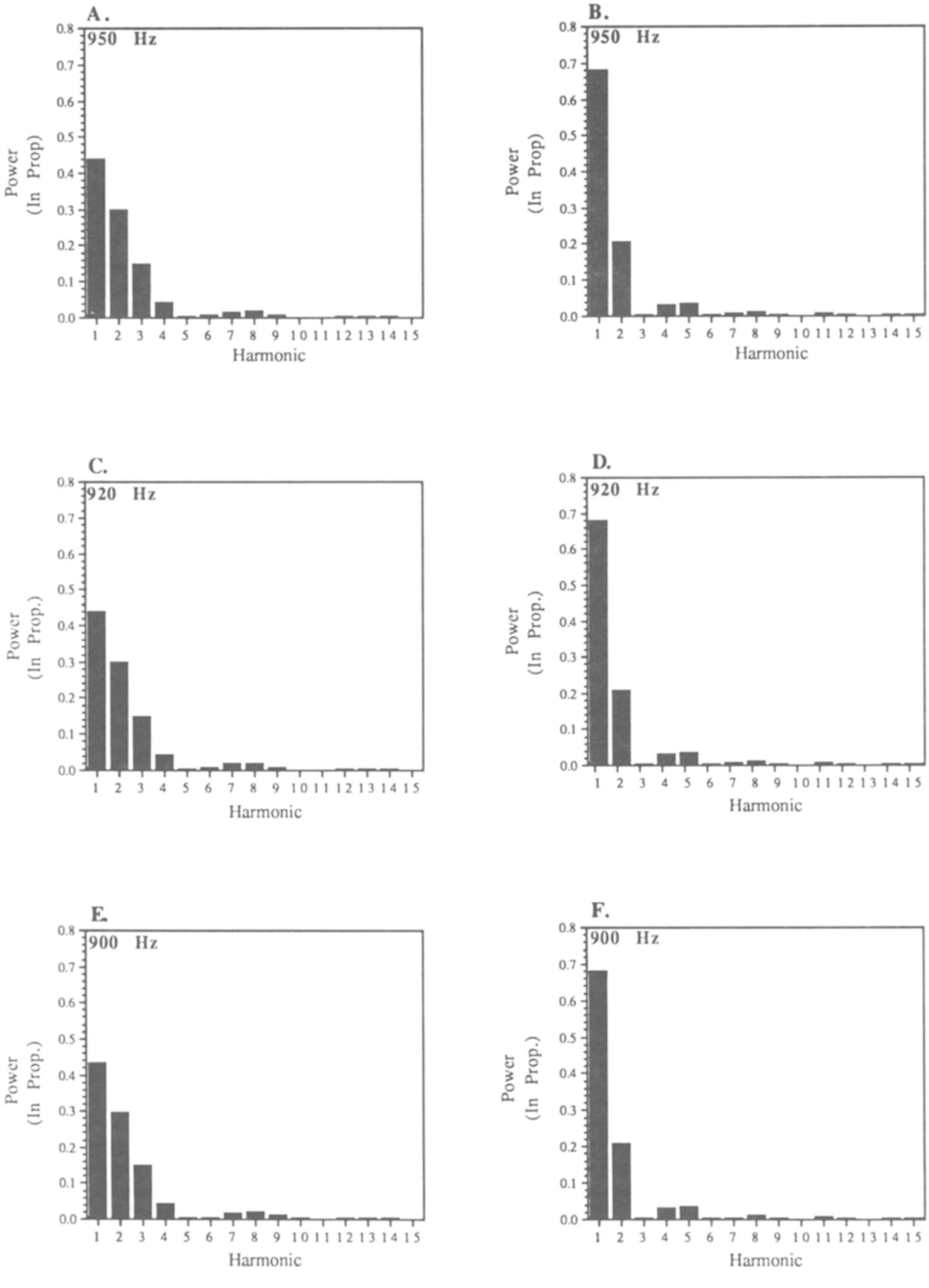

Figure 2. Power spectra (up to and including the 15th harmonic) for each of the six signals used in this study. The left side are spectra for signals having the .1878 duty cycle; the right side are spectra for signals having the .3128 duty cycle. Signals have fundamental frequencies of 950 Hz (A and B; Experiment 1), 920 Hz (C and D; Experiment 2), or 900 Hz (E and F; Experiment 2).

Table 1
Mean Reaction Times (in Milliseconds) for Judgments of Timbre and Loudness in Baseline, Filtering, and Correlated Dimensions Tasks (Experiment 1)

Task	Timbre		Loudness		Mean	
	M	Prop. Error	*M*	Prop. Error	*M*	Prop. Error
Baseline	375	.07	346	.04	360	.05
Filtering	432	.14	390	.08	411	.11
Positively correlated dimensions	332	.04	316	.03	324	.03
Negatively correlated dimensions	404	.06	364	.06	384	.06
Mean	386	.07	354	.05	370	.06

.002], because congruity effects for timbre judgments were absent in the baseline and filtering conditions.

Discussion

Experiment 1 demonstrated a clear interaction between timbre and loudness. Subjects were unable to attend selectively to either one of these dimensions; in classifying the criterial dimension, they suffered interference from variation along the irrelevant dimension.

Interestingly, the pattern of results suggests that timbre and loudness may not interact in the same way that pitch and loudness do. In particular, the two pairs of dimensions differed in how discrimination was affected by a negative correlation of dimensions. Pitch and loudness display a redundancy gain when correlated negatively (e.g., Melara & Marks, 1990c). Redundancy gain when dimensions are correlated either positively or negatively is typical of dimensions that interact *integrally*, as with the color dimensions of saturation and brightness (Garner, 1974; Garner & Felfoldy, 1970). Timbre and loudness, in contrast, produce a redundancy gain when correlated positively but a redundancy loss (i.e., interference) when correlated negatively.

This combination—gain in response to positively correlated dimensions but loss in response to negatively correlated dimensions—has also been found with the synesthetically (cross-modally) corresponding dimensions of visual color (white, black) and auditory pitch (Melara, 1989a). With color and pitch, subjects are fastest to respond to the stimulus ensembles white-high pitch and black-low pitch, resulting in (1) differences between positively and negatively correlated dimensions and (2) substantial congruity effects. Independent similarity judgments show that subjects regard white as similar to high pitch and black as similar to low pitch (Melara, 1989b). Melara (1989a, 1989b; Melara & Marks, in press; Melara & O'Brien, 1987) has argued that this and other *correspondence* interactions differ qualitatively from integral interactions because correspondence interactions are rooted in crosstalk at a *semantic* level of processing, whereas integral interactions can usually be identified with a *sensory/perceptual* level of processing. Evidence for semantic processing with corresponding dimensions includes, among other things, the fact that the pattern obtained in speeded classification is identical whether the corresponding dimensions are linguistic or nonlinguistic (Melara & Marks, in press).

Given the parallel effects of correlated dimensions on color and pitch, on the one hand, and of timbre and loudness, on the other, it is tempting to hypothesize that the latter pair do not interact as classically integral dimensions (such as pitch and loudness or saturation and brightness) do, but rather as do cross-modally corresponding ones. According to this hypothesis, crosstalk between timbre and loudness occurs at a semantic level of processing (in addition to other levels perhaps). We explore this possibility in the General Discussion.

EXPERIMENT 2
TIMBRE AND PITCH

Do timbre and pitch give rise to Garner interference in filtering tasks? And, if these dimensions do interact, do they do so in an integral manner (i.e., only show redundancy gains) or do they resemble corresponding dimensions (i.e., show a combination of gain and loss)? These are the questions asked in Experiment 2.

We also examined congruity scores. We defined congruence and incongruence in accordance with the empirically determined poles discovered earlier with pitch-

Table 2
Mean Reaction Times (in Milliseconds) to Stimuli with Congruent and Incongruent Attributes for Judgments of Timbre and Loudness in Baseline, Filtering, and Correlated Dimensions Contexts (Experiment 1)

Task Type	Timbre			Loudness		
	Stimuli		Congruity Score	Stimuli		Congruity Score
	Congruent	Incongruent		Congruent	Incongruent	
Baseline	377	376	−1	339	352	+13
Filtering	432	431	−1	372	408	+36
Correlated	332	404	+72	316	364	+48
Weighted mean	370	398	+28	336	368	+32

Note—Means are weighted to reflect the fact that congruity scores are based on fewer trials in the filtering condition than in the baseline or correlated conditions.

Table 3
Mean Reaction Times (in Milliseconds) for Judgments of Timbre and Pitch in Baseline, Filtering, and Correlated Dimensions Tasks (Experiment 2)

	Timbre		Pitch		Mean	
Task	*M*	Prop. Error	*M*	Prop. Error	*M*	Prop. Error
Baseline	345	.07	345	.06	345	.06
Filtering	407	.12	377	.10	392	.11
Positively correlated dimensions	338	.06	327	.05	333	.06
Negatively correlated dimensions	308	.04	320	.03	314	.04
Mean	350	.07	342	.06	346	.07

loudness (Melara & Marks, 1990c) and timbre-loudness (Experiment 1). Hence, high-pitched twangy sounds and low-pitched hollow sounds were defined here as congruent; low-pitched twangy sounds and high-pitched hollow sounds were defined here as incongruent. As we shall see, our scoring definitions, despite being based on these earlier experiments, were apparently different from the rules used explicitly or implicitly by our subjects.

Method

Subjects. Twenty subjects (7 men and 13 women) were paid $5 to participate.

Stimuli, Apparatus, and Procedure. The apparatus was identical to that in Experiment 1. All experimental tasks were created from a general stimulus set of four sounds. The set was formed by combining one of two possible fundamental frequencies, 920 Hz (high) and 900 Hz (low), with one of two possible duty cycles, .1878 (twangy) and .3128 (hollow). Again, rise and decay times of the "rectangular" waveforms were 10 msec. Intensity of the resulting four sounds, high-twangy, high-hollow, low-twangy, low-hollow, was held constant at 70 dB(A). The power spectra for the signals delivered to the headphones appear in panels C–F of Figure 2. The tasks and procedure were identical to those in Experiment 1.

Results

Overall performance. RTs and error rates appear in Table 3. The overall error rate was .07; RT and errors correlated .95. Baseline discriminabilities of pitch and timbre matched evenly at 345 msec.

Effects of correlated and uncorrelated dimensions. Filtering tasks (392 msec) were performed 47 msec slower than baseline tasks (345 msec), which represents a significant amount of Garner interference [$F(1,19) = 29.82$, $p < .001$]. The effect was significant for each dimension tested alone, although it was larger for the timbre judgments [62 msec; $F(1,19) = 42.11$, $p < .001$] than it was for pitch judgments [32 msec; $F(1,19) = 7.95$, $p = .01$], causing a significant task × judgment interaction [$F(1,19) = 10.60$, $p < .01$].

Correlated dimensions (324 msec) were classified 21 msec faster than baseline [$F(1,19) = 16.59$, $p < .01$], a significant redundancy gain. The gain was significant for our negatively correlated dimensions [314 msec; $F(1,19) = 26.0$, $p < .001$], but only marginally so for our positively correlated dimensions [333 msec; $F(1,19) = 4.21$, $p = .05$]. RTs to positively and negatively correlated dimensions differed significantly from each other [$F(1,19) = 13.02$, $p < .01$]. But the relatively poorer performance with positively correlated dimensions was restricted largely to timbre judgments; only here did positively correlated dimensions (338 msec) differ from negatively correlated dimensions [308 msec; $F(1,19) = 13.25$, $p < .01$].

Effects of congruity. An ANOVA of RTs showed that the stimuli we called incongruent were classified 15 msec faster overall than the stimuli we called congruent [$F(1,19) = 31.59$, $p < .001$]. The congruity scores for baseline, filtering, and correlated dimensions tasks appear in Table 4. For all three types of tasks, a significant (negative) congruity effect obtained: baseline [−12 msec; $t(19) = -5.26$, $p < .001$]; correlated dimensions [−19 msec; $t(19) = -3.55$, $p < .01$]; filtering [−15 msec; $t(19) = -3.36$, $p < .01$]. All of the congruity effects were stronger for timbre judgments than for pitch judgments, although

Table 4
Mean Reaction Times (in Milliseconds) to Stimuli with Congruent and Incongruent Attributes for Judgments of Timbre and Pitch in Baseline, Filtering, and Correlated Dimensions Contexts (Experiment 2)

	Timbre			Pitch		
	Stimuli		Congruity	Stimuli		Congruity
Task Type	Congruent	Incongruent	Score	Congruent	Incongruent	Score
Baseline	364	350	−13	364	353	−11
Filtering	427	410	−17	396	382	−14
Correlated	338	308	−30	327	320	−7
Weighted mean	366	345	−21	356	346	−10

Note—Means are weighted to reflect the fact that congruity scores are based on fewer trials in the filtering condition than in the baseline or correlated conditions.

judgment and task did not interact significantly [$F(1,19) = 1.94$, n.s.]. Only the congruity effect of pitch judgments in correlated dimensions tasks was unreliable.

Discussion

Experiment 2 demonstrated that timbre and pitch interact in speeded classification. The pattern obtained suggests an integral interaction: There was a failure of selective attention in filtering and both dimensional correlations led to redundancy gain. This pattern is reminiscent of that found with the pairs pitch and loudness (Melara & Marks, 1990c; see also Grau & Kemler-Nelson, 1988). Furthermore, the pattern is unlike that found with either timbre-loudness (Experiment 1) or pairs of synesthetically corresponding dimensions (e.g., Melara, 1989a; Melara & Marks, in press; Melara & O'Brien, 1987), for which subjects fail to display redundancy gain for one direction of dimensional correlation.

We were surprised by the direction of the present congruity effects. These effects were reasonably strong in each type of task, but they were always negative. That is, subjects responded more quickly to stimuli that we considered incongruent relative to those we called congruent. Apparently, subjects' coding of congruence-incongruence was opposite to ours: To these subjects, *hollow* corresponded to high pitch and *twangy* corresponded to low pitch.

This finding suggests that a more complex relationship of correspondences among the dimensions pitch, timbre, and loudness exists than one might have guessed a priori. For pitch and loudness, *high* corresponds to *loud*, and *low* corresponds to *soft*. For timbre and loudness, *twangy* corresponds to *loud*, and *hollow* to *soft*. These correspondences might suggest (as they did to us originally) that *high*, *loud*, and *twangy* would lie at the positive pole of their respective continua, and *low*, *soft*, and *hollow* would be at the negative pole. But the relationship that we found here between pitch and timbre indicates that the system is not simply one of polar correspondences. Rather, attributes from these three dimensions may correspond according to a two-dimensional set of relationships.[2]

GENERAL DISCUSSION

In the present study, we investigated the processes involved in classifying attributes from two pairs of auditory dimensions, timbre-loudness and timbre-pitch. When classifying values from either pair, subjects suffered Garner interference; that is, subjects were unable to attend selectively to one dimension when the other dimension was varied orthogonally. These results, coupled with our earlier findings of interaction between pitch and loudness (Melara & Marks, 1990c), suggest that attributes from all three sound dimensions are processed jointly by the perceptual system and are not separable. This conclusion is reinforced by the effects on classification found when dimensions were correlated. Perfect correlation between these auditory dimensions always altered performance, most often enhancing performance relative to baseline and, in one case (i.e., negative correlation between timbre and loudness), actually harming performance. Collectively, these findings are strong evidence that auditory timbre, pitch, and loudness are mutually interactive dimensions.

What is the mechanism of interaction among these auditory dimensions? We offer a triplex model of auditory processing: three separate and higher level auditory channels that are interconnected at one and perhaps more levels of dimensional processing. Subjects can select any particular channel, but the information accessed from the selected channel (e.g., at a decisional stage) is *contextualized* by information leaked from the other two channels. Thus, for example, when one is classifying loudnesses, perceived values on that dimension are weighted by partial output from the timbre and pitch channels. Subjects' responses are thus based on analyses of the attended dimension (loudness) pooled with analyses of the irrelevant dimensions (timbre and pitch; see Melara & Marks, 1990b, for further discussion).

What levels of processing are involved in crosstalk among auditory channels? At present, we are unable to provide a complete answer to this question, but at least some evidence is consistent with a model of perceptually early integration. Research testing classification in dual context has identified pitch and timbre as HARD dimensions and loudness as a SOFT dimension (Melara & Marks, 1990b). We found that a given auditory dimension gives consistent evidence of HARDNESS or SOFTNESS regardless of the dimension with which it is paired, as long as the two dimensions interact. These findings indicate that properties intrinsic to each individual auditory dimension (i.e., whether it is HARD or SOFT) determine how processing is affected by irrelevant variation (i.e., dual context) of an interacting auditory dimension. Such intrinsic constraints would suggest that crosstalk relations are established at a sensory/perceptual level of processing.

The full answer to the levels of processing question may depend on the particular auditory dimension or pair of dimensions tested. For example, although both the timbre-loudness and the timbre-pitch pairings were affected by correlational redundancy, only the former pair displayed sensitivity to direction of correlation (i.e., positive or negative), suggesting perhaps another level of crosstalk to determine correlational direction. The emergence of congruity effects in these experiments also suggests an interpretation in terms of later levels of cognitive processing. Earlier work with corresponding dimensions identified congruity effects within a semantic level of processing. The evidence is good that the congruity effects with corresponding dimensions derive from semantic analyses (see Melara & Marks, in press). It is conceivable that the same mechanism underlies auditory congruity effects.

We believe, however, that the case for semantic crosstalk can be made only weakly with respect to congruity effects found in the present study. Unlike literally (see, e.g., Pomerantz, 1983) or figuratively (see, e.g., Melara,

1989b) corresponding dimensions, the auditory dimensions tested here were not based firmly in semantic match versus mismatch. As noted before, we began with arbitrary decisions about which stimuli were congruent and which were incongruent. Thus, it is possible that the significant congruity effects obtained had nothing to do with "congruity" per se, but were based in the relatively greater (and serendipitous) perceptual salience of those stimuli that we called "congruent." Without an identifiable basis in the meaningful correspondence of attributes, it is hazardous to assume semantic processing simply because a congruity effect has been obtained (see Melara & Marks, 1990c, for elaboration on this point).

We conclude, therefore, that although auditory interactions may have their locus in levels other than sensory/perceptual processes, at present the pattern is too ambiguous to identify what those other levels might be. Nonetheless, in the present study we have convincingly established that auditory channels do crosstalk. Particularly important was the finding that auditory timbre yields Garner interference in intraclass context established by either pitch or loudness. To the best of our knowledge, this is the first complete test of the effects of single context on timbre classification (but see Iverson & Krumhansl, 1989, for a partial analysis consistent with the results reported here).

These findings are important theoretically, because they establish a link between the effects on timbre of single and dual context. In another investigation (Melara & Marks, 1990b), we have shown that timbre classifications suffer no interference in dual context when paired with either pitch or loudness. But those findings were ambiguous because absence of interference might be due to dimensional separability. In the present study, however, we have shown that spectral timbre is not separable from either pitch or loudness. In so doing, we have disambiguated the other findings and confirmed that timbre is a HARD interacting dimension. In our laboratory, we are working to extend these conclusions—that is, to identify the auditory processes that characterize and distinguish HARD from SOFT interactions.

REFERENCES

ASHBY, F. G., & TOWNSEND, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, **93**, 154-179.

CLARK, H. H., & BROWNELL, H. H. (1975). Judging up and down. *Journal of Experimental Psychology: Human Perception & Performance*, **1**, 339-352.

CLARK, H. H., & BROWNELL, H. H. (1976). Position, direction, and their perceptual integrality. *Perception & Psychophysics*, **19**, 328-334.

DEVALOIS, R. L., & DEVALOIS, K. K. (1987). *Spatial vision*. New York: Oxford University Press.

DURLACH, N. I., BRAIDA, L. D., & ITO, Y. (1986). Toward a model for discrimination of broadband signals. *Journal of the Acoustical Society of America*, **80**, 63-72.

GARNER, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.

GARNER, W. R. (1981). The analysis of unanalyzed perceptions. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 119-139). Hillsdale, NJ: Erlbaum.

GARNER, W. R., & FELFOLDY, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, **1**, 225-241.

GRAHAM, N. (1981). Psychophysics of spatial frequency channels. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 1-26). Hillsdale, NJ: Erlbaum.

GRAU, J. W., & KEMLER-NELSON, D. G. (1988). The distinction between integral and separable dimensions: Evidence for the integrality of pitch and loudness. *Journal of Experimental Psychology: General*, **117**, 347-370.

GREEN, D. M. (1988). *Profile analysis: Auditory intensity discrimination*. New York: Oxford University Press.

HIRSCH, J., HYLTON, R., & GRAHAM, N. (1982). Simultaneous recognition of two spatial-frequency components. *Vision Research*, **22**, 365-375.

IVERSON, P., & KRUMHANSL, C. (1989). Pitch and timbre interaction in isolated tones and in sequences. *Journal of the Acoustical Society of America*, **86** (Suppl. 1), S58.

LEPAGE, E. L. (1987a). Frequency-dependent self-induced bias of the basilar membrane and its potential for controlling sensitivity and tuning in the mammalian cochlea. *Journal of the Acoustical Society of America*, **82**, 139-154.

LEPAGE, E. L. (1987b). A spatial template for the shape of tuning curves in the mammalian cochlea. *Journal of the Acoustical Society of America*, **82**, 155-164.

LOCKHEAD, G. R. (1966). Effects of dimensional redundancy on visual discrimination. *Journal of Experimental Psychology*, **72**, 95-104.

LOCKHEAD, G. R. (1972). Processing dimensional stimuli: A note. *Psychological Review*, **79**, 410-419.

LOCKHEAD, G. R. (1979). Holistic versus analytic process models: A reply. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 746-755.

MARKS, L. E. (1987). On cross-modal similarity: Auditory-visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 384-394.

MARKS, L. E., HAMMEAL, R. J., & BORNSTEIN, M. H. (1987). Perceiving similarity and comprehending metaphor. *Monographs of the Society for Research in Child Development*, **52**(1, Serial No. 215).

MELARA, R. D. (1989a). Dimensional interaction between color and pitch. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 69-79.

MELARA, R. D. (1989b). Similarity relations among synesthetic stimuli and their attributes. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 212-231.

MELARA, R. D., & MARKS, L. E. (1990a). Dimensional interactions in language processing: Investigating directions and levels of crosstalk. *Journal of Experimental Psychology: Learning, Memory, & Cognition*.

MELARA, R. D., & MARKS, L. E. (1990b). HARD and SOFT interacting dimensions: Differential effects of dual context on classification. *Perception & Psychophysics*, **47**, 307-325.

MELARA, R. D., & MARKS, L. E. (1990c). Perceptual primacy of dimensions: Support for a model of dimensional interaction. *Journal of Experimental Psychology: Human Perception & Performance*.

MELARA, R. D., & MARKS, L. E. (in press). Processes underlying dimensional interactions: Correspondences between linguistic and nonlinguistic dimensions. *Memory & Cognition*.

MELARA, R. D., & O'BRIEN, T. P. (1987). Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology: General*, **116**, 323-336.

POMERANTZ, J. R. (1983). Global and local precedence: Selective attention in form and motion perception. *Journal of Experimental Psychology: General*, **112**, 515-540.

POMERANTZ, J. R. (1986). Visual form perception: An overview. In E. C. Schwab & H. C. Nusbaum (Eds.), *Pattern recognition by humans and machines: Visual perception* (pp. 1-30). New York: Academic Press.

POMERANTZ, J. R., PRISTACH, E. A., & CARSON, C. E. (1989). Attention and object perception. In B. E. Shepp & S. Ballesteros (Eds.), *Object perception: Structure and process* (pp. 53-89). Hillsdale, NJ: Erlbaum.

POSNER, M. I. (1964). Information reduction in the analysis of sequential tasks. *Psychological Review*, **71**, 491-504.

PRICE, J. M. (1979). Software timing for 6500 series microcomputers. *Behavior Research Methods & Instrumentation*, **11**, 568-571.

WALKER, P., & SMITH, S. (1984). Stroop interference based on the synesthetic qualities of auditory pitch. *Perception*, **13**, 75-81.

WALKER, P., & SMITH, S. (1986). The basis of Stroop interference involving the multimodal correlates of auditory pitch. *Perception*, **15**, 491-496.

WOOD, C. C. (1975). Auditory and phonetic levels of processing in speech perception: Neurophysiological and information-processing analyses. *Journal of Experimental Psychology: Human Perception & Performance*, **1**, 3-20.

ZWICKER, E. (1986). A hardware cochlear nonlinear preprocessing model with active feedback. *Journal of the Acoustical Society of America*, **80**, 146-153.

NOTES

1. The higher level channels proposed here are thought to be distinct from physiologically based channels suggested, for example, in the processing of spatial and temporal frequency (see, e.g., DeValois & DeValois, 1987; Graham, 1981; Hirsch, Hylton, & Graham, 1982; LePage, 1987a, 1987b; Zwicker, 1986). Nonetheless, we do not rule out the possibility that neural channels may give rise later in processing to channels that are specific to psychologically meaningful (i.e., primary) dimensions.

2. An analogous two-dimensional structure has been reported previously among the dimensions auditory pitch, auditory loudness, and visual size (Marks, Hammeal, & Bornstein, 1987). Subjects judge "high pitch" to be similar to "small" and "loud." However, "loud" is judged to be more similar to "large" than to "small." It remains open whether the present results and these earlier results have a similar basis in processing.

(Manuscript received July 24, 1989;
revision accepted for publication March 20, 1990.)