Perception & Psychophysics
2002, 64 (2), 198-207

Influence of tonal context and timbral variation on perception of pitch

CATHERINE M. WARRIER and ROBERT J. ZATORRE
McGill University and Montreal Neurological Institute, Montreal, Quebec, Canada

In this study, spectral timbre's effect on pitch perception is examined in varying contexts. In two experiments, subjects detected pitch deviations of tones differing in brightness in an isolated context in which they compared two tones, in a tone-series context in which they judged whether the last tone of a simple sequence was in or out of tune, and in a melodic context in which they determined whether the last note of familiar melodies was in or out of tune. Timbre influenced pitch judgments in all the conditions, but increasing tonal context allowed the subjects to extract pitch information more accurately. This appears to be due to two factors: (1) The presence of extra tones creates a stronger reference point from which to judge pitch, and (2) the melodies' tonal structure gives more cues that facilitate pitch extraction, even in the face of conflicting spectral information.

Pitch and timbre are two of the building blocks of music. Variations in pitch lead to a melodic line, whereas variations of timbre are usually heard as different instrumentations. Do pitch and timbre interact? Listening to an orchestra, one can hear a continuous melody being played when different instruments switch off playing separate parts of this melody, even if each note of the melody is played by a different instrument, a compositional style called klangfarbenmelodie or hocket. A more striking demonstration of this phenomenon occurs with sung melodies, in which the changing vocal timbres associated with speech do not alter perception of the melody. This implies that one can follow the fundamental frequency ($F0$) of a series of tones, even when their spectral shapes differ, which argues for the separability of pitch and timbre. However, since both pitch and spectral timbre are rooted in the frequency dimension of sound, it should not be surprising if they interact under some circumstances. To test this idea, one could look at people's perceptions of pitch and/or timbre when both the $F0$ and the spectral shape of the tones differ. This paper investigates the interaction between pitch and timbre, focusing on spectral timbre's influence on pitch perception as a function of context.

The literature in which interactions between pitch and timbre are examined has yielded contradictory results. Some researchers have found that the timbre of a tone affects its perceived pitch (e.g., Krumhansl & Iverson, 1992, Experiment 1; Melara & Marks, 1990a, 1990b, 1990c; Platt & Racine, 1985; Singh & Hirsh, 1992; Wapnick & Freeman, 1980), whereas others have found no effect of timbre on pitch perception (e.g., Krumhansl & Iverson, 1992, Experiments 2 and 3; Semal & Demany 1991, 1993). It seems that those studies presenting tones in the absence of other tones tend to find an interaction between pitch and timbre, whereas studies presenting tones within the context of other tones find no such interaction (but see also Demany & Semal, 1993, in which pitch and timbre difference thresholds for isolated tones were not affected by variation in the irrelevant dimension).

A same–different paradigm was used by Singh and Hirsh (1992) to determine the perceived pitch of isolated residue tones, tones having no component at $F0$. Six timbres were synthesized, each containing four consecutive harmonics, the lowest of which could be the second, third, fourth, or up to the seventh harmonic. Each pair of tones could differ in $F0$, spectral composition, or both. Subjects indicated whether timbre was the same or different, and whether pitch stayed the same, went up, or went down. When the harmonic numbers and $F0$ moved in the same direction, subjects correctly reported the direction of pitch change. However, when harmonic number and $F0$ moved in opposite directions, this created a conflict. If the change in $F0$ was 4% or greater, the direction of $F0$ change dominated pitch judgments. However, when the change in $F0$ was less than 4%, the direction of harmonic change dominated pitch judgments. Therefore, pitch and timbre were found to be separable only with a change in $F0$ of 4% or higher.

Platt and Racine (1985) used a tuning paradigm to examine differences in the perceived pitch of pure and complex tones. Larger tuning deviations were made when a complex tone was tuned to a pure tone than when both the standard and the test tones were pure tones, suggesting an interaction between pitch and timbre. When single tones are classified into categories, another paradigm performed without a tonal context, pitch and timbre are also found to interact (Krumhansl & Iverson, 1992, Experiment 1; Melara

This research was supported by funding from the Medical Research Council of Canada (Operating Grant MT 11541). We thank Pierre Ahad for his technical assistance, as well as Al Bregman for use of his laboratory facilities. Correspondence concerning this article should be addressed to C. M. Warrier, Communication Sciences and Disorders, Francis Searle Building, Northwestern University, Evanston, IL 60208 (e-mail: cwarrier@northwestern.edu).

& Marks, 1990a, 1990b, 1990c). Melara and Marks performed a series of experiments in which pitch, timbre, and loudness interactions were investigated, using the Garner (1974) speeded classification method; they observed how redundant and interfering information from the unattended dimension affected reaction times, and looked at classification reaction times for selective and divided attention to these dimensions with varying orientations of stimulus axes. These studies illustrate a consistent interaction between pitch and timbre: Correlating information from the unattended dimension enhances classification, whereas competing information disrupts it (Melara & Marks, 1990a, 1990b, 1990c). Krumhansl and Iverson (1992) replicated the Garner classification results with different stimuli, finding pitch and timbre to interact in isolation. They then went on to test for these interactions within longer sequences and found a much different result: Changes in pitch of the context tones affected perception of the pitch, but not of the timbre, of the test tones, and changes in timbre of the context tones weakly affected the perception of the timbre, but not of pitch.

This lack of interaction is seen in other studies examining pitch and timbre in a tonal context. Semal and Demany (1991, 1993) performed a series of experiments, finding no interaction between pitch and timbre. In both studies, they presented eight-tone sequences and asked subjects to judge whether the last tone was identical to or different from the first tone. The six interpolated tones varied by $F0$, spectral timbre, amplitude envelope, and intensity. However, only the $F0$ of the interference tones affected performance; timbral variations had no effect. Therefore, it was argued that pitch and timbre are completely separable. In a study using a similar paradigm, variations in pitch were not found to significantly affect timbre perception (Starr & Pitt, 1997).

It appears that interactions between pitch and timbre tend to occur in situations in which test tones are presented in the absence of other tones. Pitch and timbre are found to influence each other when methodologies that present tones in the absence of a tonal context are used, such as comparing a single tone with another single tone by tuning (Platt & Racine, 1985), making qualitative judgments (Sing & Hirsh, 1992; Wapnick & Freeman, 1980), or classifying single tones into categories (Krumhansl & Iverson, 1992, Experiment 1; Melara & Marks, 1990a, 1990b, 1990c). In contrast, no evidence of an interaction is found by those studies that present tones in the context of other tones, as when interpolating tones are manipulated (Semal & Demany, 1991, 1993; Starr & Pitt, 1997) or the fourth tones of two seven-tone sequences are compared, a modified interpolated tone paradigm (Krumhansl & Iverson, 1992, Experiments 2 and 3).

The fact that tonal context can affect pitch perception has been well established (Deutsch, 1972a, 1972b, 1982a; Deutsch & Roll, 1974; Dewar, Cuddy, & Mewhort, 1977; Krumhansl, 1979; Krumhansl & Castellano, 1983). After performing a series of studies on the matter, Krumhansl concluded that "the representation of pitch consists of a pattern of interrelationships that is highly specific to the tonal system of the musical context. Further, this pattern of interrelationships, once established, has implications for the processing of subsequent musical events" (1979, p. 372). The present study tests whether or not tonal context's effect on pitch can also affect the interaction between pitch and timbre, possibly explaining the differences seen in previous studies. Our goal was to examine the effect of context specifically, while controlling for effects of interpolating tones by utilizing a methodology that avoided the use of any interpolating tones. Two contexts were used in Experiment 1: (1) an isolated context in which only two tones were presented for each trial, and (2) a melodic context in which these same tones were presented as the final tone of a familiar melody. The task in each condition was to make pitch judgments while ignoring changes in timbre. We hypothesized that the presence of other tones in the melodic context would provide more of a tonal reference point from which to judge pitch, so that people would make more accurate pitch judgments and be better able to ignore differences in spectral composition in the melodic context than in the isolated context.

EXPERIMENT 1

Method

Subjects

Eleven McGill undergraduates participated in this experiment. None had any extensive musical training (average training = 1.25 years). Normal hearing was determined by self-report. All were compensated for their time and attention.

Stimuli

The stimuli were digitally synthesized on an IBM-compatible 386 computer using MITSYN software (Henke, 1981). All test tones were 500 msec in duration, with rise and fall times of 10 msec. Contextual tones used in the melodic condition varied in duration according to the score. Three timbres were created by varying the relative intensities of 11 harmonics, keeping the intensity of the fundamental frequency constant. One emphasized the lower harmonics (low), another emphasized the middle harmonics (middle) and sounded brighter than low, and the last emphasized the higher harmonics (high) and sounded brighter than middle (see Figure 1). $F0$ was kept at a constant level in each timbre in order to avoid basing differences between tones on $F0$ strength. Individual tones ranged in fundamental frequency from 164.81 to 1108.70 Hz. (E_3 to $C\sharp_6$), with test tones always presented between 261.63 (C_4) and 480.35 Hz ($A\sharp_4 + 52¢$), and all were equated for loudness. The stimuli were presented with MAPLE (Manager of Auditory Perception and Linguistic Experiments) software (Achim, Ahad, & Bregman, 1992) and run through a passive Tchebyshev function low-pass filter having a 3-dB cutoff at 8000 Hz and a slope of –142 dB per octave. Sounds were presented binaurally at 75 dB SPL over Seinheisser HD424 headphones in a soundproof booth.

Procedure

Isolated context. The concepts of pitch and timbre were explained, with timbre being explained as a change in *instrument* or *sound quality*, and exemplars of the three spectral shapes were presented at a common $F0$. Each trial consisted of two tones presented with an interstimulus interval of 100 msec. The first tone was randomly presented at one of six frequencies corresponding to whole tone steps starting at C_4 (261.63 Hz) and continuing to $A\sharp_4$

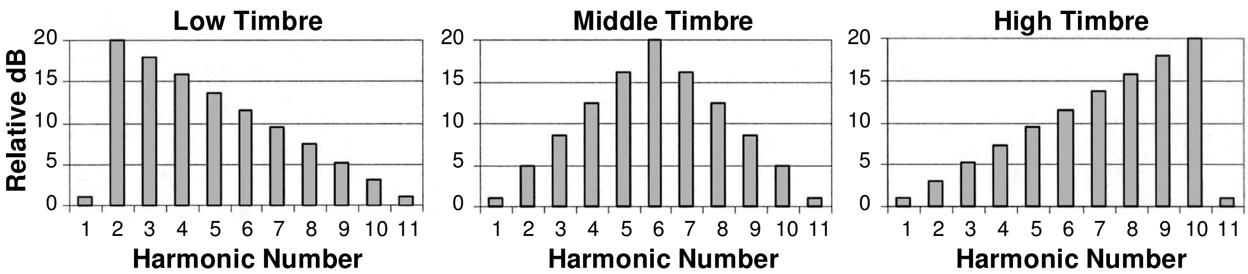

Figure 1. Spectral shapes of timbres. Differences in the intensity level of each harmonic are shown relative to the other harmonics.

(466.16 Hz). The second tone was 0 (same), 17¢, 35¢, or 52¢ higher than the first tone, corresponding to the 1%, 2%, and 3% differences at which Singh and Hirsh (1992) found pitch and timbre to interact. (Cents are logarithmically equal steps in the frequency dimension, each semitone being 100¢ apart.) Each pair of test tones could be presented in one of three timbre pairings: (1) same (low–low, high–high), (2) small difference (low–middle, high–middle), and (3) large difference (high–low, low–high). Each pairing was presented at all levels of $F0$ difference. The subjects were informed that they would be making pitch judgments on these tones and that they should ignore any timbre changes as much as possible. When the tones were presented, the subjects were instructed to indicate whether the pitch of the second tone was the same as or different from the first tone by pressing a key on the computer keyboard. If they responded *different*, they were asked to indicate how different on a scale of 1–3 where 1 = *slightly different* and 3 = *very different*. The subjects completed 36 trials at their own pace after completing a block of 6 practice trials.

Melodic context. For the melodic context, the task was to determine whether the last note of a melody was in tune or not. One of two melodies, familiar to all the subjects, was presented on each trial: "Oh, Susanna" or "The Blue Danube Waltz" (see Figure 2). These melodies were chosen for their different endings; the four final notes of "Blue Danube" all share the same pitch, whereas "Oh, Susanna" ends in a downward scalar motion. Each melody, excepting the last note, was presented in tune in one of three keys, so that the last note should end on one of six whole steps from C_4 to $A\sharp_4$, and was presented in one of the three timbres described above.[1] The last note could continue with the same spectral shape as the melody or sound in one of the other two, thus creating the same three timbre pairs as the isolated condition, and could be 0¢ (same), 17¢, 35¢, or 52¢ sharper than the correct final note. After each melody, the subjects were instructed to indicate whether the last note of the melody was in or out of tune, ignoring any timbre differences. If they responded *out of tune*, they were asked to indicate how much out of tune on a scale of 1–3, where 1 = *slightly out of tune* and 3 = *very out of tune*. The subjects completed 36 trials of each melody at their own pace, after a block of 6 practice trials. We predicted that if there was a difference in responses between the two melodies, the subjects would do better on "Blue Danube" trials, because they could compare the pitch of the last note directly with that of the penultimate note.

Results

The results for each trial were coded on a scale of 0–3, where 0 indicated a *same* response and 1–3 corresponded to the scales of 1–3 described above. Analyses of variance (ANOVAs) were performed within subjects as repeated measures tests with three factors: context (isolated vs.

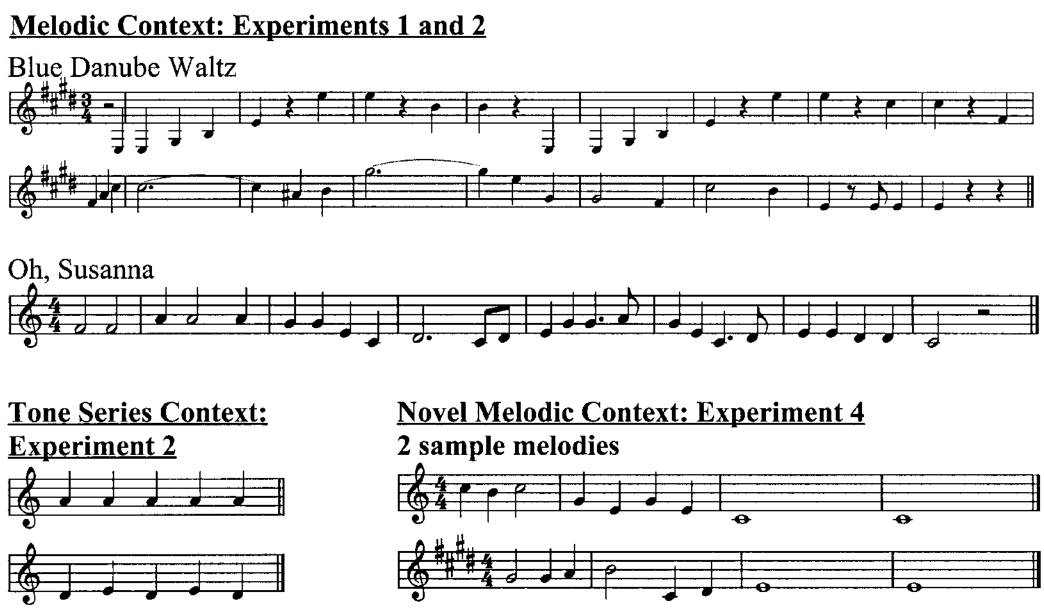

Figure 2. Stimulus contexts shown in musical notation.

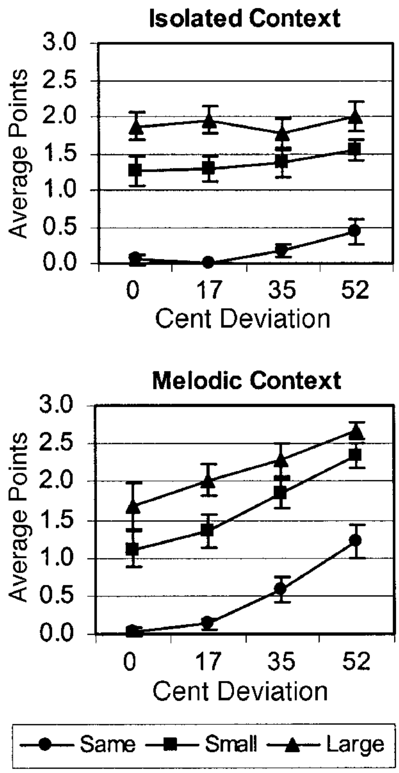

Figure 3. Experiment 1 results averaged across subjects, scored on a scale of 0–3, where 0 means *same pitch* (see the text for details). In the timbre differences legend, "Same" denotes same timbre, "Small" denotes a small difference in timbre, and "Large" denotes a large difference in timbre. Standard error bars are shown.

melodic), $F0$ (0¢, 17¢, 35¢, or 52¢), and timbre-pair type (same, small difference, or large difference). Tukey's HSD test (Q statistic) was used for post hoc analyses.

In the melodic condition, responses to both melodies were similar, with increasing responses to larger timbre differences and larger $F0$ deviations (see Figure 3). An interaction between timbre-pair type and melody was found, so we analyzed responses to each timbre-pair type across melodies. In this analysis, $F0$ deviations in different-timbre trials in "Blue Danube" were judged larger than in "Oh, Susanna" [$F(1,10) = 17.21$, $p < .01$, for small timbre difference, and $F(1,10) = 10.11$, $p = .01$, for large timbre difference]. When we collapsed across timbre-pair types, the subjects judged $F0$ differences consistently across melodies. Since both melodies differed significantly from the isolated context in the same direction and manner, while differing much less between themselves, data from the two melodies were averaged together and called the *melodic condition* in further analyses.

Responses to the isolated condition also increased with increasing differences in timbre. But unlike the melodic condition, responses did not increase reliably with increasing $F0$ deviations (see Figure 3).

Statistical comparisons between the two contexts showed a main effect of timbre-pair type [$F(2,20) = 131.22$, $p < .001$] and an interaction between context and $F0$ [$F(3,30) = 19.68$, $p < .001$]. In investigating this interaction, we conducted two types of post hoc analyses. First, we compared the ratings of each $F0$ level between contexts, collapsing across timbre pairs. No differences were found between contexts in either the 0¢ or the 17¢ deviation [$F(1,10) = 1.04$ and 1.47, respectively; $p > .2$ for both]. The subjects rated 35¢ and 52¢ trials as more different/out of tune in the melodic than in the isolated context, indicating better discriminability of $F0$ in the melodic condition [$Q(8,30) = 7.7$, $p < .05$, for 35¢; $Q(8,30) = 12.4$, $p < .05$, for 52¢].

We also investigated which of the $F0$ differences were detected within each context by comparing ratings on *different-F0* trials with those of the *same-F0* trials, again collapsing over timbre-pair types. The only $F0$ difference detected in the isolated condition was the 52¢ deviation [$Q(4,30) = 5.22$, $p < .05$]. In contrast, both the 35¢ and 52¢ deviations were detected in the melodic condition, again arguing for better $F0$ discrimination in the melodic condition [$Q(4,30) = 7.71$, $p < .05$, for 35¢; $Q(4,30) = 13.88$, $p < .05$, for 52¢].

Discussion

The most salient effect seen in this experiment was the influence spectral shape had on the detection of small changes in $F0$ in both the isolated and the melodic conditions. At all levels of $F0$, tones that differed in timbre were judged to be more different in pitch than when timbre was the same. This replicates the results found in Singh and Hirsh's (1992) study, in which timbre differences interfered with detection of $F0$ deviations of a similar magnitude. However, this experiment also showed that placing tones within a melodic context allowed listeners to perceive $F0$ differences that were undetectable without this contextual information. This facilitation occurred despite the continuing strong effect of timbre on pitch judgments.

This facilitation of pitch discrimination within a melodic context could be due to the fact that the melodies are creating a tonal reference point to which the test tones can be related. According to Deutsch (1982b), "in listening to sequences, we process not only the individual tones, but also the melodic intervals between them. These intervals then provide a framework of pitch relationships to which the test tones can be anchored" (p. 302). The tonal structure of the melodic condition, unavailable in the isolated condition, helps us to focus in on the pitch information carried in the tones and thus seems to facilitate the separation of timbre from the pitch percept.

Although addressing some of the issues regarding the interactions between pitch, timbre, and context, Experiment 1 also raises new questions. For example, exactly what aspect of the melodic context allows pitch to be perceived more accurately than when this context is not present? Would the simple presence of other tones be enough to show this effect, or is the tonal structure of the melodies necessary?

A second experiment therefore included a third context in which the same tones were placed at the end of nonmelodic five-tone sequences that mimicked the endings of the melodies used in the melodic condition of Experiment 1. We predicted that if the effect of context seen in Experiment 1 was caused by the simple presence of other tones, the results of the tone series condition would be no different than the melodic condition. However, if the tonal structure of the notes is necessary to see an effect of context, the tone series condition would yield results more like the isolated condition. If the effect was due to both the tones' presence and their tonal structure, performance on the tone series condition would lie between the isolated and the melodic conditions, owing to the presence of one element, but not the other.

EXPERIMENT 2

Method

Subjects and Stimuli

Twelve McGill undergraduates participated in this experiment. None had any extensive musical training (average length = 1.67 years). No subject reported any hearing loss. All were compensated for their time and attention. The stimuli were synthesized and presented as in Experiment 1.

Procedure

The isolated and melodic contexts were presented as in Experiment 1, with the exception that the 17¢ deviation of $F0$ was dropped because it was indistinguishable from the 0¢ difference in both conditions. In addition to these two contexts, Experiment 2 contained a third context in which the test tones were placed at the end of two simple five-tone sequences. This context was presented in the same manner as the melodic condition, except that instead of hearing melodies, the subjects heard tone sequences. One sequence simply repeated one tone five times, whereas the other alternated between one tone and a whole tone above it, mirroring the endings heard in the melodic condition (Figure 2). The order of the three conditions was counterbalanced across subjects to control for order effects.

Results

As in Experiment 1, the results for each trial were coded on a scale of 0–3. ANOVAs were performed within subjects as repeated measures tests with three factors: context (isolated, tone series, and melodic), $F0$ (0¢, 35¢, and 52¢), and timbre-pair type (same, small difference, and large difference). Tukey HSD tests were performed for post hoc analyses.

The pattern of results in both the isolated and the melodic contexts replicated the results of Experiment 1 (see Figure 4). Responses to both melodies again increased with increases in both timbre and $F0$ difference. $F0$ deviations in different-timbre trials were again judged larger than in "Oh, Susanna" [$F(1,11) = 16.02$, $p < .01$, for small spectral differences; $F(1,11) = 23.08$, $p = .001$, for large spectral differences], but the differences between melodies were much smaller than their mutual differences from the isolated condition, and so the two melodies were grouped together for further analysis as the melodic condition. No consistent differences were found between responses in the alternating- and the repeating-tone conditions, so they were also averaged together for further analysis as the tone series condition [$F(1,11) = 3.91$, $p > .05$].

Statistical comparisons between contexts uncovered a three-way interaction between context, timbre-pair type, and $F0$ [$F(8,88) = 3.83$, $p < .01$]. In investigating this interaction, we again conducted two types of post hoc analyses, both using Tukey's HSD test. First, we compared the ratings of each $F0$ level between contexts at each of the three spectral pairings. The pattern of results indicated an increased discriminability of $F0$ with increasing levels of context.

Most of the differences found were seen when comparing the isolated with the melodic context. When spectral shape did not differ, both the 35¢ and the 52¢ deviations were judged larger in the melodic context [$Q(9,88) = 6.74$, $p < .05$, for 35¢; $Q(9,88) = 9.38$, $p < .05$, for 52¢]. With

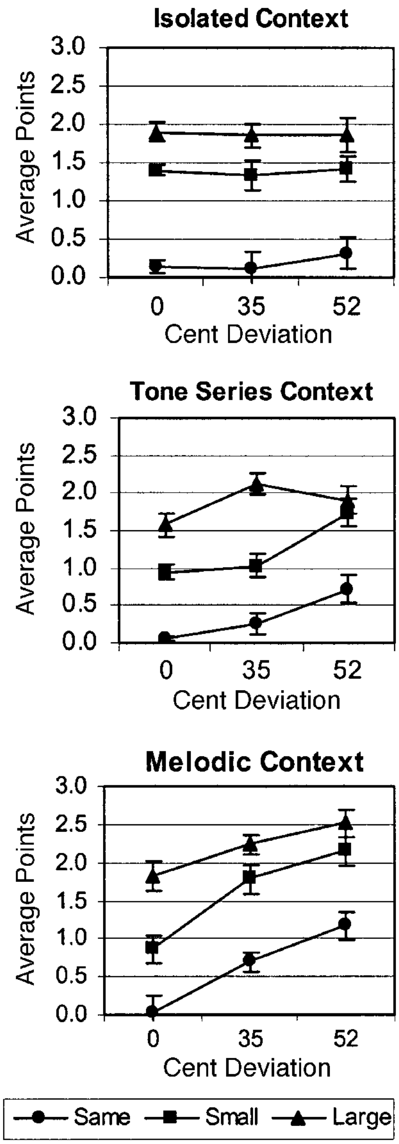

Figure 4. Experiment 2 results averaged across subjects, scored on a scale of 0–3, where 0 means *same pitch* (see the text for details). In the timbre differences legend, "Same" denotes same timbre, "Small" denotes a small difference in timbre, and "Large" denotes a large difference in timbre. Standard error bars are shown.

small spectral differences, 0¢ deviations were judged smaller, and 35¢ and 52¢ deviations were judged larger in the melodic condition [$Q(9,88) = 6.01$, $p < .05$, for 0¢; $Q(9,88) = 5.29$, $p < .05$, for 35¢; $Q(9,88) = 8.78$, $p < .05$, for 52¢]. With large spectral differences, the 52¢ deviation was judged larger in the melodic context [$Q(9,88) = 7.34$, $p < .05$].

Because the tone series results fall between those of the isolated and the melodic contexts, they differ only slightly from either one. Comparing the tone series and melodic contexts, with small spectral differences, the 35¢ deviation was judged larger in the melodic context [$Q(9,88) = 8.78$, $p < .05$]. With large spectral differences, the 52¢ deviation was judged larger in the melodic condition [$Q(9,88) = 6.74$, $p < .05$]. The only difference seen between the isolated and the tone series contexts was that people judged same-$F0$ trials for the small timbre difference as being smaller in the tone series context [$Q(9,88) = 5.29$, $p < .05$].

We also investigated which $F0$ differences were detected by comparing different-$F0$ trial ratings to same-$F0$ trial ratings within each timbre difference for each context—for example, within large timbre difference trials, was the 35¢ deviation judged different from same-$F0$ trials. None of the $F0$ differences were detected in the isolated condition. Three of the six $F0$ deviations were detected in the tone series condition. The 52¢ difference was detected in both the same and the small spectral difference trials [$Q(9,88) = 7.70$, $p < .05$, for same; $Q(9,88) = 9.26$, $p <.05$, for small], and the 35¢ deviation was detected in the large spectral difference trials [$Q(9,88) = 6.37$, $p < .05$]. All $F0$ deviations were detected in the melodic condition. This analysis also indicates increasing discriminability of $F0$ with increasing levels of context.

Discussion

The isolated and melodic conditions in Experiment 2 replicated Experiment 1. The effect of spectral shape on pitch perception was again the most salient effect, being highly significant in all three contexts. Thus, the subjects were more likely to judge pitch as more different when timbre differed than when it was the same, regardless of differences in $F0$. When comparing the isolated and the melodic conditions, we saw the same pattern of results as those found in Experiment 1—namely, despite timbre's influence on pitch perception, $F0$ differences that were undetectable in the isolated context were perceived when heard within the context of a melody.

The pattern of results in both post hoc analyses indicates an increasing ability to extract pitch information from tones with conflicting spectral information as the level of context increases from isolated to tone series to melodic context. This result was most striking when looking at which $F0$ differences were reliably detected, relative to no $F0$ change. By this analysis, none of the $F0$ changes in the isolated condition were discriminated, whereas all were discriminated in the melodic condition. The tone series context fell halfway between them, since three out of six $F0$ differences were reliably detected, thus supporting our third prediction. This result was predicted if the facilitation of pitch discrimination seen in the melodic condition was due to both the presence of other tones and the tonal structure of those tones.

Another difference between the tone series and the melodic conditions is the length of the sequences. The five-tone sequences used in the tone series condition were shorter in length and number of tones than the melodic sequences. Experiment 3 addresses this potential confound by comparing tone series contexts differing only in length: a five-tone series and a longer series matching the length of the melodic context. We predicted no difference in the results from these two lengths of tone series.

EXPERIMENT 3

Method

Subjects

Five graduate students participated in this experiment. The average length of musical training was 2.6 years. Normal hearing was determined by self-report.

Stimuli

This experiment used two variations in spectral shape, the high and low timbres described in Experiment 1 and two $F0$ deviations: 0¢ (same) and 52¢. The stimuli were created and presented through the same computer system as that in Experiment 1 in a sound-attenuate d room.

Procedure

Two variations of the tone series context described in Experiment 2 were used in this experiment. Context 1 (short sequence) used the 5-tone alternating and repeating sequences exactly as presented in Experiment 2. Context 2 (long sequence) extended these series to 34 tones, matching the average length of the melodies in Experiments 1 and 2, 8.75 sec. Each series was presented so that the last note should end on one of six whole steps from C_4 to $A\#_4$. The last note of each sequence could be played correctly—that is, at the correct $F0$ with the same spectral shape as the previous note, or in a different spectral shape and/or $F0$ than the correct final note. This resulted in four different stimulus types: same-pitch–same-timbre, same-pitch–different-timbre, different-pitch–same-timbre, and different-pitch–different-timbre. The subjects were asked to determine whether the last note was in tune or out of tune with the rest of the sequence. Eight practice trials were presented before the test conditions, each containing 96 trials (24 trials of each stimulus type).

Results and Discussion

No differences were seen between responses to the short sequences and the long sequences [$F(1,4) = 0.04$, $p > .8$]. No significant interaction between sequence length and stimulus type was found [$F(3,12) = 0.55$, $p > .6$; see Figure 5].

Since no difference was found between responses to short tone sequences of the length used in Experiment 2 and those to tone sequences matched in length to the melodies used in Experiment 2, the differences seen between the tone series and melodic conditions of Experiment 2 cannot be attributed to a difference in length.

EXPERIMENT 4

The last experiment in this study addressed the possibility that differences in response instructions could be influencing the results. Although the task in each context is

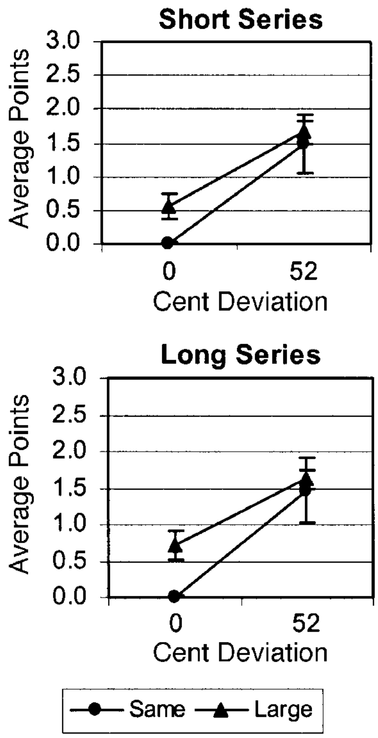

Figure 5. Experiment 3 results averaged across subjects (each subject performed both series), scored on a scale of 0–3 where 0 means *same pitch* (see the text for details). In the timbre differences legend, "Same" denotes same timbre, and "Large" denotes a large difference in timbre. Standard error bars are shown.

to detect deviations in pitch, the semantics of asking this question changes between contexts. One responds same versus different in the isolated context, whereas one responds in tune versus out of tune in the melodic and tone series contexts. This experiment addresses this issue and, by using novel melodies, investigates the role that the familiarity of the melodies played in the results seen in Experiments 1 and 2. Did the results seen in these previous experiments rely heavily on the fact that familiar melodies were heard, or would the "syntax" of tonality generate a strong enough reference point to obtain the same pattern of results with novel melodies? We predicted that the strict tonality of our novel melodies would similarly create a reference point from which to judge pitch, as with the familiar melodies in Experiments 1 and 2, and that the results from this experiment would follow the same pattern.

Method

Subjects

Ten graduate students participated in this experiment, 5 of whom also participated in Experiment 3. None had any extensive musical training (average training = 1.6 years). Normal hearing was determined by self-report.

Stimuli

This experiment used two variations in spectral shape, the high and low timbres described in Experiment 1 and two $F0$ deviations, 0¢ (same) and 52¢. Stimuli were created and presented through the same computer system as in Experiment 1 in a sound-attenuated room.

Procedure

On each trial, the subjects were presented with one of four short melodies. When played correctly, the last two notes of each melody had the same pitch (see Figure 2 for examples). The last note of each melody could be played correctly—that is, with the same spectral shape at the same $F0$ as the previous note—or in a different spectral shape and/or $F0$, as in Experiment 3. The melodies were constructed following strict rules of tonality and did not present the pitch of the test tones before the end of the melody. When played correctly, each could end on one of six whole steps from C_4 to $A\sharp_4$. The subjects were divided into two groups, each with slightly different instructions. Group 1 was asked to determine whether the last note of each melody was played in tune or out of tune, disregarding any change in timbre. Group 2 was asked to determine whether the last two notes of each melody had the same or a different pitch, also disregarding any change in timbre. Both groups were presented with identical stimuli, the only difference between groups being the phrasing of the task requirements. Eight practice trials were presented before the 144 test trials, 36 of each stimulus type.

Results

No significant differences were found between the two subject groups or between melodies [$F(1,8) = 1.01, p > .3$, and $F(11,99) = 1.31, p > .2$, respectively]. The results were consistent with the previous melodic conditions of Experiments 1 and 2 (see Figure 6)—namely, a main effect of pitch [$F(1,9) = 36.04, p < .001$] and a main effect of timbre [$F(1,9) = 17.17, p < .005$] were found. Post hoc examination of which $F0$ differences were reliably detected, done by comparing the 0¢ and 52¢ ratings within each timbre pair, using Wilcoxon's signed rank test, showed that both 52¢ deviations were detected ($Z = -2.803, p = .005$). This pattern is consistent with the melodic conditions of Experiments 1 and 2.

Discussion

No differences were seen between the two subject groups, those who were asked to judge whether the tones at the end of melodies were played either at the same or a different pitch, and those who judged whether they were played in or out of tune. Therefore, the differences seen between conditions in Experiments 1 and 2 cannot be attributed to the phrasing of the task instructions. This does not necessarily imply that people are performing the task in the same way in the isolated and the melodic conditions. Even though the physical difference upon which each judgment is made is the same in both conditions, there is more information available in the melodic condition. The in/out of tune question is meant to encourage listeners to make better use of the contextual cues that are unavailable in the isolated condition, in which an in/out of tune judgment is irrelevant. It is possible that listeners might be able to ignore the context in other cases. However, as is seen in this experiment, even when instructed to compare the last two tones of each melody, thus rendering the rest of the melody irrelevant to the task, the listeners apparently did not ignore the context, since they were able to reliably detect the $F0$ differences present in the stimuli. This

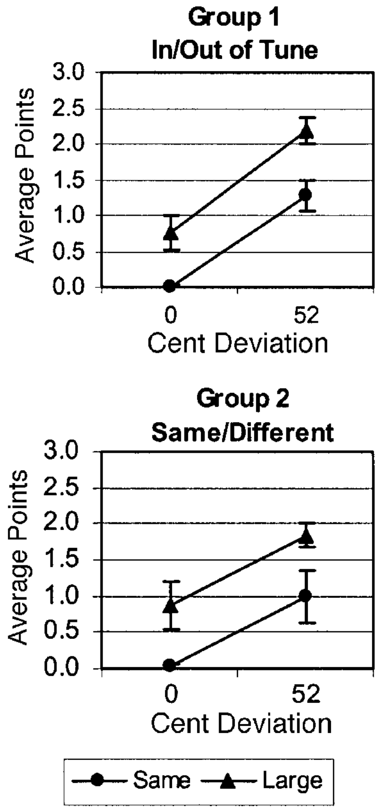

Figure 6. Experiment 4 results averaged across subjects (each subject participated under one set of instructions only), scored on a scale of 0–3, where 0 means *same pitch* (see the text for details). In the timbre differences legend, "Same" denotes same timbre, and "Large" denotes a large difference in timbre. Standard error bars are shown.

contrasts with results from the isolated conditions in Experiments 1 and 2, in which $F0$ differences were undetectable.

As was predicted, the results obtained with novel tonal melodies mimicked the pattern of results seen with familiar melodies in Experiments 1 and 2. In those experiments, we found the presence of familiar melodies to enhance pitch perception in the face of spectral discrepancies between the test tones. We attributed this result to both the presence and the tonal structure of the tones making up the melodic context. Experiment 4 extends this result to novel melodies, assuring us that it was not merely the familiarity of the melodies that facilitated task performance in the melodic conditions of Experiments 1 and 2. The similar results seen with both familiar and novel melodies highlight the robustness of this effect.

GENERAL DISCUSSION

In this study, the interaction of pitch and timbre was investigated by examining the extent of the influence of timbre on pitch perception under three different contextual conditions: isolated, tone series, and melodic. At the small differences in $F0$ examined in this study, spectral shape always influenced pitch discrimination. In all the conditions we examined, subjects judged the pitch of different-timbre trials as more different than same-timbre trials regardless of differences in $F0$. Experiments 1 and 2 showed that increasing tonal context facilitates pitch extraction, despite timbre's effect on pitch perception. Experiments 3 and 4 confirmed that contextual effects were not due to confounding factors, such as different lengths of stimuli in the melodic and tone series conditions or differences in instruction semantics.

The fact that spectral shape had an influence on pitch judgments indicates that pitch and timbre do interact and are not completely separable dimensions of sound. This result conflicts with the studies using interpolating tones, which find pitch and timbre to be completely separable. However, the interpolated tone paradigm differs from ours in a couple of crucial aspects. It requires the subject to remember a standard tone over a period of time usually lasting a few seconds, whereas our paradigm allows a more temporally direct comparison. Also, in the interpolated tone studies, test tones were always presented with the same spectral shape. It was the interpolated tones that differed in timbre and were meant to be ignored by the subjects. In our study, the test tones themselves differed in spectral shape; that is, the test tone could vary in both pitch and timbre from the comparison/correct tone. Ignoring the irrelevant property of timbre in the test tones may be a harder task than ignoring it in irrelevant tones. Indeed, this difference in methodology covaries in the pitch and timbre literature with whether or not test tones are heard in the context of other tones. Perhaps the reason we do not see a complete separation of pitch and timbre in the melodic condition is because we varied both attributes within our test tones.

At larger differences in $F0$, spectral shape is not likely to affect pitch judgments. In Singh and Hirsh's (1992) study, which used a paradigm analogous to our isolated condition, pitch and timbre did not interfere with each other at $F0$ differences of 4% and higher. This 4% difference corresponds approximately to a 70¢ difference, which is larger than our largest $F0$ deviation of 52¢. Therefore, on the basis of this previous study, we predict that increasing $F0$ in any of the contexts used in the present study should decrease spectral shape's effect on pitch judgments.

The low level of musical training of our subjects could also be affecting the size of spectral shape's effect in this study. A speeded classification study comparing musicians and nonmusicians on pitch and timbre perception showed that nonmusicians' pitch judgments are more affected by timbre differences than are musicians' (Pitt, 1994). Trained musicians have more experience not only in making pitch judgments, but also in determining whether different instruments are in tune with each other, which involves discriminating pitch while ignoring differences in spectral shape. This practice would probably enable musicians to outperform nonmusicians on our paradigm. However, pilot testing in our laboratory leads us to expect musicians to show the same pattern of results across conditions as the nonmusicians—namely, that differences in spectral shape would interfere with their pitch judgments less with increasing tonal context. More research needs to be done to confirm this expectation.

Tonal context appears to be playing an important role in this study. Placing a tone within a tonal context seems to create a better point of reference from which other tones can be judged. Since our Western tonal system is based on ratios, without a reference point listeners cannot determine whether or not a single note is in tune or not (unless they have absolute pitch). A two-tone interval has a specific ratio between $F0$s, which can be in tune or out of tune. Hearing more notes creates more of a tonal reference point by which one can judge whether each tone is in or out of tune with the rest of the notes present (e.g., Deutsch 1972b; Dewar et al., 1977; Krumhansl, 1979). Crowder (1993) states that when listening to a sequence of tones, "even when there is no obvious melody, individual tones are not heard independently of a tonal context" (p. 137). In addition to supporting this model, our results suggest that a tonal reference point can lessen timbre's influence on pitch perception, allowing people to extract pitch information even more efficiently, and that increasing levels of tonal context allow more accurate pitch extraction.

The melodic conditions in this study showed the greatest number of facilitating effects. But can we be sure that this is due to their tonal structure, and not simply to their higher structural complexity in comparison with the other contexts? One way to test this idea would be to include a random context that has a complex structure, arguably more complex than the simple melodies used in this study, and is atonal. Therefore, if structural complexity is the facilitating factor, task performance should be higher in the random context, and if tonal structure in particular is the facilitating factor, performance should be better in the melodic context. We made this comparison in another study using a similar task and found that performance with a random context falls between the isolated and the melodic contexts (Warrier, Belin, Merlet, & Zatorre, 1999). This result argues against the idea that the facilitation we see is due solely to the complexity of the context and supports the idea that the rich tonal structure of the melodies is facilitating pitch extraction.

Another factor that may be contributing to the differences seen between contexts in this study is the preparatory effect of expectation. Expectations of final notes were created by the tone series, owing to their repetitive nature, and by the novel melodies through their strong tonal structure. The familiar melodies, in addition to their tonal structure, created expectancies of the final note simply by being familiar to the subjects; they knew from memory what the final tone should sound like. Expectation causes a mental priming to occur that focuses attention to a particular $F0$. Bharucha (1994) suggests that this priming involves the formation of a mental image before the last tone is heard, with which the physical final tone can be compared.

One method of testing whether expectancy is playing a role is to test whether a decrease in performance is seen when the stimulus is unexpected. In a study (Bharucha & Stoekig, 1986) in which the expectancy of chords was looked at, chords were presented at the end of a progression, and listeners were asked to decide whether that last chord was in tune or out of tune. Performance on unexpected chords was lower than that on expected chords. In the present study, the timbre of the last note could be considered expected or unexpected in this way. Hearing a melody or series of tones being played in a specific timbre sets up the expectation that the last note will be played in the same timbre. So, the expected timbre in this study was that in which the beginning of the trial was presented (same-timbre trials). When the timbre of the last note differed from the beginning timbre, this was unexpected from that trial's context (different-timbre trials). Our listeners were more accurate when the last note was presented in the expected timbre. Evidence such as this suggests the tonal context leading up to the test tones not only creates a tonal framework from which to judge pitch, but also creates expectancies in the mind of the listener that are either fulfilled or violated.

These concepts, melodic context creating a tonal framework and expectation playing a preparatory role in the melodic condition, both help to explain the differences found between the isolated and the melodic conditions in this study but do not explain the differences found between the two melodies. Although both "Blue Danube" and "Oh, Susanna" create a tonal framework helpful in discriminating pitch and generate strong expectancies as to what the final note should be, people were slightly more affected by timbre in "Blue Danube" trials. We predicted any difference between the melodies to be in the opposite direction, since "Blue Danube" contains the extra cue of being able to compare the pitch of the final note with those directly preceding it. However, no differences were seen between the repeating and alternating tone series contexts in Experiment 2, indicating that the melodies' endings were not the cause of the differences seen here.

One possible reason for the slightly better results seen in "Oh, Susanna" comes from the smaller interval sizes than those found in "Blue Danube." As Deutsch (1982b) explains, "there is considerable evidence that melodic sequences are processed more effectively when these are composed of smaller size rather than larger (reflecting the operation of the principle of Proximity)" (p. 302; see also Deutsch, 1978). More evidence is needed to exactly determine the reason a difference between the two melodies was found. However, it is important to remember that although the difference between the two melodies sheds more insight into the nature of the overall interaction between pitch and timbre, this difference is quite small, as compared with their mutual differences from the isolated condition.

Although we have been describing our results as showing that timbre's influence on pitch perception lessens with increasing tonal context, this could also be interpreted as an increase in the salience of the pitch dimension, making it easier to ignore spectral variations. These interpretations are not mutually exclusive. Our paradigm does not allow us to distinguish between the two interpretations, but both imply an interaction between pitch and timbre.

In conclusion, this study has demonstrated a progressive improvement in pitch discrimination with three increasing levels of context. The extra tones heard in each trial of the tone series condition, as compared with the isolated condition, created a stronger reference point from which to judge the test tone, thus enhancing pitch discrimination. The facilitation seen in the melodic condition, as compared with the tone series condition, appears to be due not to the increase in number of tones, but to the structured tonality of the melodies. Their rich internal structure lends more cues as to the tonality of the melody than do simple repeating or alternating sequences. The auditory system takes advantage of this extra information when extracting pitch, both with and without conflicting spectral information.

REFERENCES

Achim, A., Ahad, P., & Bregman, A. (1992). *MAPLE: Manager of auditory perception and linguistic experiments* [Computer program]. Montreal: McGill Speech and Hearing Lab.

Bharucha, J. J. (1994). Tonality and expectation. In R. Aiello & J. Sloboda (Eds.), *Musical perceptions* (pp. 213-239). New York: Oxford University Press.

Bharucha, J. J., & Stoekig, K. (1986). Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception & Performance,* **12**, 403-410.

Crowder, R. G. (1993). Auditory memory. In S. McAdams (Ed.), *Thinking in sound* (pp. 113-145). Oxford: Oxford University Press.

Demany, L., & Semal, C. (1993). Pitch versus brightness of timbre: Detecting combined shifts in fundamental and formant frequency. *Music Perception,* **11**, 1-14.

Deutsch, D. (1972a). Effect of repetition of standard and comparison tones on recognition memory for pitch. *Journal of Experimental Psychology,* **93**, 156-162.

Deutsch, D. (1972b). Mapping of interactions in the pitch memory store. *Science,* **175**, 1020-1022.

Deutsch, D. (1978). Delayed pitch comparisons and the principle of proximity. *Perception & Psychophysics,* **23**, 227-230.

Deutsch, D. (1982a). The influence of melodic context on pitch recognition judgment. *Perception & Psychophysics,* **31**, 407-410.

Deutsch, D. (1982b). The processing of pitch combinations. In D. Deutsch (Ed.), *The psychology of music* (pp. 271-316). New York: Academic Presss.

Deutsch, D., & Roll, P. L. (1974). Error patterns in delayed pitch comparison as a function of relational context. *Journal of Experimental Psychology,* **103**, 1027-1034.

Dewar, K. M., Cuddy, L. L., & Mewhort, D. J. (1977). Recognition memory for single tones with and without context. *Journal of Experimental Psychology: Human Learning & Memory,* **3**, 60-67.

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

Henke, W. L. (1981). *MITSYN: A coherent family of command-level utilities for time signal processing* [Computer program]. Belmont, MA: Author.

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology,* **11**, 346-374.

Krumhansl, C. L., & Castellano, M. A. (1983). Dynamic processes in music perception. *Memory & Cognition,* **11**, 325-334.

Krumhansl, C. L., & Iverson, P. (1992). Perceptual interactions between musical pitch and timbre. *Journal of Experimental Psychology: Human Perception & Performance,* **18**, 739-751.

Melara, R. D., & Marks, L. E. (1990a). Hard and soft interacting dimensions: Differential effects of dual context on classification. *Perception & Psychophysics,* **47**, 307-325.

Melara, R. D., & Marks, L. E. (1990b). Interaction among auditory dimensions: Timbre, pitch, and loudness. *Perception & Psychophysics,* **48**, 169-178.

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,* **16**, 398-414.

Pitt, M. A. (1994). Perception of pitch and timbre by musically trained and untrained listeners. *Journal of Experimental Psychology: Human Perception & Performance,* **20**, 976-986.

Platt, J. R., & Racine, R. J. (1985). Effect of frequency, timbre, experience, and feedback on musical tuning skills. *Perception & Psychophysics,* **38**, 543-553.

Semal, C., & Demany, L. (1991). Dissociation of pitch from timbre in auditory short-term memory. *Journal of the Acoustical Society of America,* **89**, 2404-2410.

Semal, C., & Demany, L. (1993). Further evidence for an autonomous processing of pitch in auditory short-term memory. *Journal of the Acoustical Society of America,* **94**, 1315-1322.

Singh, P. G., & Hirsh, I. J. (1992). Influence of spectral locus and *F*0 changes on the pitch and timbre of complex tones. *Journal of the Acoustical Society of America,* **92**, 2650-2661.

Starr, G., & Pitt, M. A. (1997). Interference effects in short-term memory for timbre. *Journal of the Acoustical Society of America,* **102**, 486-494.

Wapnick, J., & Freeman, P. (1980). Effects of dark–bright timbral variation on the perception of flatness and sharpness. *Journal of Research in Music Education,* **28**, 176-184.

Warrier, C., Belin, P., Merlet, I., & Zatorre, R. J. (1999). fMRI study examining effect of melodic context on pitch discrimination. *Society for Neuroscience Abstracts,* **25**, 1629.

NOTE

1. Our filter cutoff was at 8000 Hz, with a –142-dB slope per octave. All of the harmonics of the test tones were well within the filter (our highest test tone at 466 Hz has an 11th harmonic of 5126 Hz). The highest harmonics of the highest *contextual* tones, however, lie above this cutoff frequency and are therefore presented at a decreased decibel level. This may cause a slight change in the timbre of those few contextual tones. For example, at 9000 Hz there is a 20 dB drop, and at 10000 Hz there is a 38-dB drop. So, the highest three harmonics of our highest contextual tone are most likely inaudible to most ears. However, the frequency range of these highest harmonics is quite high, and many adults would have difficulty hearing tones in this range. In addition, the tones affected are rare in our melodies, being present only in "Blue Danube's" highest jumps. So we do not feel that the slight change in timbre of these rare tones caused by our filter has any real effect on our results.

(Manuscript received July 12, 1999;
revision accepted for publication April 13, 2001.)