A Rule of Thumb: The Bandwidth for Timbre Invariance Is One Octave

Music Perception
Summer 2001, Vol. 19, No. 1, 121–126

A Rule of Thumb: The Bandwidth for Timbre Invariance Is One Octave

STEPHEN HANDEL & MOLLY L. ERICKSON
University of Tennessee, Knoxville

Listeners were unable to determine whether two different notes separated by an octave or more were played by the identical or a different wind instrument and similarly were unable to determine whether a vowel sung at different pitches separated by an octave or more was sung by the identical or a different soprano or mezzo-soprano.

Received August 26, 2000; accepted April 26, 2001.

THE term *timbre* is used in two seemingly contradictory ways. First, technically the timbre of an instrument or voice is defined at one pitch and loudness (ANSI, 1973). Thus, the timbre of each note will change as a function of F_0, as the spectral centroids and onset transients vary across pitch. Second, timbre refers to a perceptual quality that characterizes an instrument or voice across its range and is assumed ultimately to be due to the mechanics of sound production (Hourdin, Charbonneau, & Moussa,1997a, 1997b; Kendall, Carterette, & Hajda, 1998). It is the latter conceptualization of timbre as an invariant quality based on perceivable transformations across pitch and/or loudness that is assumed to underlie the ability to identify one instrument or voice. To provide a visual analogy, although facial pictures of babies and adults are obviously different (corresponding to timbres at different F_0s), Pittinger, Shaw, and Mark (1979) have shown that people can match baby pictures to adult pictures on the basis of an understanding of normal growth curves.

We would argue that to reconcile the uses of the term timbre, there must be some form of acoustic transformation perceived by the listener that provides a way to predict the quality of one note on the basis of the quality of another note played (or sung) by the same instrument (or same singer).

Address correspondence to Stephen Handel, Department of Psychology, University of Tennessee, Knoxville, TN 37996-0900. (e-mail: shandel@utk.edu)

ISSN: 0730-7829. Send requests for permission to reprint to Rights and Permissions, University of California Press, 2000 Center St., Ste. 303, Berkeley, CA 94704-1223.

Here, we will operationally define the bandwidth of the perceptual transformations, termed *timbre invariance,* by the listener's ability to determine whether or not two notes at different pitches were played (or sung) by the identical instrument (or singer). We argue that for wind instruments and classically trained singing voices, this invariance is limited to roughly one octave. Thus, there are overlapping octave bands of equivalent timbre.

Experiment 1: Wind Instruments

The purpose of Experiment 1 was to determine how well listeners could determine whether two instrumental notes at different pitches were played on identical or different instruments.

METHOD

All 22 participants were musically naïve undergraduates at the University of Tennessee who received course credit.

The six instruments used were the clarinet, English horn, French horn, oboe, trombone, and trumpet. The notes ranged from G_3 to C_6. All stimuli were recorded from the McGill University Master Samples Compact Disks (Opolko & Wapnick, 1989) at 44.1 kHz (16-bit resolution) The stereo format on the compact disc was converted to monophonic by summing the left and right channels. Because the notes on the compact disks vary greatly in loudness and duration, all tones were normalized to have roughly the same amplitude and the same duration, thus eliminating these two variables as potential cues. In order to preserve onset cues, each stimulus consisted of the first 1.5 to 2.5 s of the tone. Clicks associated with truncated offsets were eliminated by applying a uniform smoothing function. Several studies, beginning with those by Berger (1964) and Saldanha and Corsa (1964), showed that offset transients did not affect instrument identification (although the onset transient did).

On every trial, the two different notes were presented in an ABABA sequence: the lower pitch was always the initial note, and there was a 0.5-s silence between each pair of notes. All identical and all different instrument pairings are shown in Table 1. The note pair C_4/G_4 was constructed for all six instruments, creating six identical pairs and 30 (6 x 5) different pairs. The most dissimilar note pair, G_3/C_6, was constructed for the clarinet and trumpet. Here, there would be four pairs: Clarinet G_3/Clarinet C_6, Trumpet G_3/Trumpet C_6, Clarinet G_3/ Trumpet C_6, Trumpet G_3/Clarinet C_6. A total of 35 identical instrument pairings and 78 different instrument pairings were used. The 113 trials were broken into four blocks and presented in counterbalanced order across subjects.

The subjects responded on a four-point scale ranging from "(1) very certain identical instrument" to "(4) very certain different instruments." Before the experimental trials began, subjects were presented examples using stringed instruments to familiarize them with the procedure. The tones were presented at a comfortable listening level, approximately 65 dB (SPL).

RESULTS

Preliminary results using Tukey's Honestly Significant Different Test on the average judgments indicated that the note pairs with less than one octave separation and note pairs with more than one octave separations could

TABLE 1
Pairs of Notes Used in Experiment 1

Category	Notes	Instruments	Number of Identical Pairs	Number of Different Pairs
Less than one octave	C_3/G_3	French horn	1	0
	G_3/C_4	English horn, French horn, trombone	3	0
	C_4/G_4	Clarinet, English horn, French horn, oboe, trumpet, trombone	6	30
	G_4/C_5	Clarinet, oboe, trombone	3	0
	C_5/G_5	English horn, oboe	2	0
One octave	G_3/G_4	Clarinet, English horn, French horn, trumpet, trombone	5	20
	C_4/C_5	Clarinet, English horn, French horn, oboe	4	12
More than one octave	C_3/C_5	French horn, trombone	2	2
	G_3/G_5	Clarinet, English horn, trumpet	3	6
	G_3/C_6	Clarinet, trumpet	2	2
	C_4/C_6	Clarinet, oboe, trumpet	3	6
	G_4/C_6	Oboe	1	0

be collapsed into single categories. The boxplots for the resulting six categories are shown in Figure 1. Outliers that were misjudged are labeled.

The most important outcome is that the one-octave separation delimits the range at which listeners were able to determine whether two notes were played by the identical instrument or by different instruments. Beyond one octave, listeners could not discriminate between the two possibilities and perceived all but one pair of notes to have been played by different instruments. At less than one octave, listeners often made the reverse error and incorrectly perceived pairs of notes to have been played by the same instrument. The average rating for 15 of the 30 pairs of notes played by different instruments was less than 2.5.

It is important to note that, with two exceptions, the judgments for pairs of different instruments separated by one or more octaves were quite similar (e.g., Trumpet G_3/Clarinet C_6 = 3.78 and Clarinet G_3/Trumpet C_6 = 3.23). Across the 22 combinations of instruments (see Table 1), the average absolute difference was 0.41. The two exceptions both involved the French horn and trombone: (1) French horn C_3/Trombone C_5 = 2.57 whereas Trombone C_3/French horn C_5 = 3.7 and (2) French horn G_3/Trombone G_4 = 1.66 whereas Trombone G_3/French horn G_4 = 2.95. We have no explanation for this asymmetry.

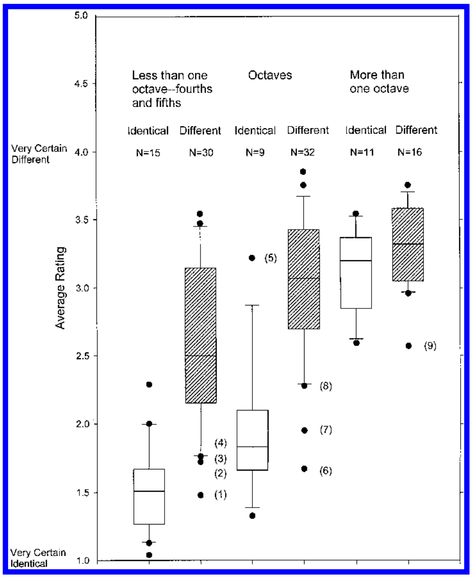

Fig. 1. Experiment 1: Woodwind instruments. Average rating of whether the two notes were played by the identical or different instruments. "N" refers to the number of different exemplars of each category. Instrument pairs that were mislabeled: (1) French horn C_4/ Trombone G_4; (2) Trombone C_4/Trumpet G_4; (3) Trumpet C_4/Oboe G_4; (4) English horn C_4/Oboe G_4; (5) Clarinet G_3/Clarinet G_4; (6) French horn G_3/Trombone G_4; (7) Oboe C_4/English horn C_5; (8) Clarinet C_4/ Oboe C_5; (9) French horn C_3/Trombone C_5.

Experiment 2: Classically Trained Singing Voices

To study the ability of experienced (e.g., choral directors with Master's degrees in voice) and inexperienced listeners to identify classically trained singers at different pitches, Erickson, Perry, and Handel (in press) made use of an oddball paradigm.

METHODS

Stimuli consisted of sung "ahs" produced by two sopranos and two mezzo-sopranos. From the original recordings of at least 3 s of sustained ah, a 1-s sample was obtained from the midpoint so as to obtain samples with steady vibrato rates. Shaping functions were applied to onsets and offsets. The recorded samples were of approximately the same loudness. Trials of three stimuli were constructed where two of the three stimuli in each trial were produced by the same singer at two different pitches and the third stimulus was produced by a different singer. For each singer, three same-singer conditions were constructed: (1) A_3/A_5 separated by two octaves, (2) C_4/F_5 separated by an octave and a fourth, and (3) G_4/B_4 separated by a major third. For each singer and each condition, the "odd" stimulus was varied across the three remaining singers and across the pitches A_4, C_4, G_4, B_4, F_5, and A_5. This design created 216 trials based on 4 singers × 3 conditions × 3 singer-pairs × 6 pitches. For each trial, the order of the three stimuli was randomized. The resulting 216 trials were presented in random order to 12 experienced listeners and 19 inexperienced listeners.

Listeners could repeat stimuli as many times as necessary. Listeners indicated which of the three stimuli they perceived to be the different singer by clicking the icon for that stimulus. One question of interest in this study was, how do the same-singer pitch separation conditions affect listeners' ability to discriminate which is the different singer? A second question of interest was whether or not discrimination accuracy varied depending on whether the two singers were from the same voice category (e.g., soprano/soprano) or from different voice categories (e.g., soprano/mezzo-soprano).

RESULTS

The percentages of correct responses as a function of the pitch separation, the subject's experience, and voice category are shown in Figure 2. Identification of the oddball singer was excellent if the note separation of the first singer was a major third, but identification dropped nearly to chance if the note separation was more than an octave (C_4/F_5 and A_3/A_5). Experienced listeners were significantly better than inexperienced listeners only at the major third separation for both the different voice category trials [$t(29) = 2.9$, $p < .005$] and same voice category trials [$t(29)=4.0$, $p<.001$]. Moreover, identification was always significantly better if the oddball singer was in a different voice category for both experienced and inexperienced listeners [$t(11) = 5.1$, $p < .001$, and $t(18) = 6.18$, $p < .001$, respectively].

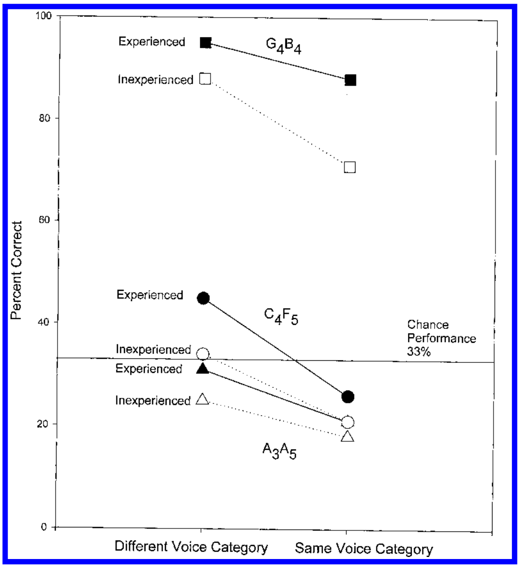

Fig. 2. Experiment 2: Classically trained singing voices. Percentage of correct responses for identifying the different oddball singer.

Discussion

For a set of wind instruments, subjects tended to judge two notes separated by more than one octave as being played by different instruments and tended to judge two notes separated by fourths and fifths as being played by the same instrument. Although there may be a bias for consonant intervals to be perceived as being played by the same instrument, listeners still judged that two notes were played by different instruments for the highly consonant one-and two-octave intervals. In fact, the maximum discrimination between same and different instruments occurred at the one-octave separation. In similar fashion, subjects could not judge which two of three notes were sung by the same singer if the notes were separated by more than one octave even for the consonant A_3/A_5 interval. The latter result was found for both experienced and inexperienced listeners, suggesting that the results found for the wind instruments were not solely due to the choice of inexperienced subjects. These outcomes strongly imply that listeners can extrapolate the timbre of an instrument or voice over only a relatively short pitch range and that the invariance of instrumental and voice timbre is based on similar perceptual principles. We would argue that the tasks used here are the simplest possible, so that the estimate of a one-octave timbre "window" is probably an upper bound.[1]

References

ANSI (1973). *Psychoacoustical terminology. S3.20*. New York: American National Standards Institute.

Berger, K. W. (1964). Some factors in the recognition of timbre. *Journal of the Acoustical Society of America*, *36*(10), 1888–1891.

Erickson, M. L., Perry, S., & Handel, S. (in press). Discrimination functions: can they be used to classify singing voices? *Journal of Voice*.

Hourdin, C., Charbonneau, G., & Moussa, T. (1997a). A multidimensional scaling analysis of musical instruments time-varying spectra. *Computer Music Journal*, *21*, 40–55.

Hourdin, C., Charbonneau, G., & Moussa, T. (1997b). A sound synthesis technique based on multidimensional scaling of spectra. *Computer Music Journal*, *21*, 56–67.

Kendall, R. A., Carterette, E. C., & Hajda, J. M. (1998). Perceptual and acoustical features of natural and synthetic orchestral instrument tones. *Music Perception*, *16*, 327–364.

Opolko, F., & Wapnick, J. (1989). *McGill University master samples user's manual*. Montreal: McGill University, Faculty of Music.

Pittinger, J. B., Shaw, R. E., & Mark, L. S. (1979). Perceptual information for the age levels of faces as a higher-order invariant of growth. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 478–493.

Saldanha, E. L., & Corso, J. F. (1964). Timbre cues and the identification of musical instruments. *Journal of the Acoustical Society of America*, *36*(11), 2021–2026.

1. We thank John Gratz, who prepared the instrumental notes and ran Experiment 1.