

Perception of dissonance by people with normal hearing and sensorineural hearing loss^{a)}

Jennifer B. Tufts,^{b)} Michelle R. Molis, and Marjorie R. Leek

Army Audiology and Speech Center, Walter Reed Army Medical Center, 6900 Georgia Avenue NW, Washington, DC 20307

(Received 21 July 2004; revised 21 April 2005; accepted 4 May 2005)

The purpose of this study was to determine whether the perceived sensory dissonance of pairs of pure tones (PT dyads) or pairs of harmonic complex tones (HC dyads) is altered due to sensorineural hearing loss. Four normal-hearing (NH) and four hearing-impaired (HI) listeners judged the sensory dissonance of PT dyads geometrically centered at 500 and 2000 Hz, and of HC dyads with fundamental frequencies geometrically centered at 500 Hz. The frequency separation of the members of the dyads varied from 0 Hz to just over an octave. In addition, frequency selectivity was assessed at 500 and 2000 Hz for each listener. Maximum dissonance was perceived at frequency separations smaller than the auditory filter bandwidth for both groups of listeners, but maximum dissonance for HI listeners occurred at a greater proportion of their bandwidths at 500 Hz than at 2000 Hz. Further, their auditory filter bandwidths at 500 Hz were significantly wider than those of the NH listeners. For both the PT and HC dyads, curves displaying dissonance as a function of frequency separation were more compressed for the HI listeners, possibly reflecting less contrast between their perceptions of consonance and dissonance compared with the NH listeners. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1942347]

PACS number(s): 43.66.Jh, 43.66.Sr, 43.75.Cd [RAL]

Pages: 955–967

I. INTRODUCTION

Research and anecdotal evidence suggest that many listeners with sensorineural hearing loss (SNHL) do not perceive music normally, even with the use of hearing aids and/or cochlear implants (e.g., Gfeller *et al.*, 2000; Chasin, 2003). An understanding of how SNHL affects the perception of music, besides providing greater insight into the nature of SNHL, may suggest approaches to improving hearing aids and cochlear implants for both music and speech listening.

Several aspects of the musical signal are affected by SNHL. deLaat and Plomp (1985) found that people with SNHL had greater difficulty recognizing a melody presented simultaneously with two other melodies than did normal-hearing people. Other reports have demonstrated poor complex pitch perception and inaccuracies in pure-tone octave judgments (Arehart and Burns, 1999), abnormal pitch-intensity shifts, frequency difference limens, and pitch-matching variability (Burns and Turner, 1986), and reduced pitch strength (Leek and Summers, 2001) in people with SNHL. To the extent that pitch perception and the identification and segregation of melodic lines are important in music listening, deficits in these areas may alter the perception of music by hearing-impaired listeners.

Another important aspect of musical expression that may be affected by SNHL is the perception of consonance

and dissonance. Consonance and dissonance are perceptual attributes of musical intervals that convey variation in musical tension. Generally speaking, a musical interval is described as consonant if it sounds harmonious and restful, while an interval is described as dissonant if it sounds discordant and tense. Music theorists have classified the thirteen musical intervals of the Western tradition in terms of consonance and dissonance. The so-called “perfect” intervals—the unison, octave, fifth, and fourth—are considered to be highly consonant. The major and minor thirds and sixths are considered to be imperfect consonances, that is, less consonant than the perfect intervals; the major and minor seconds and sevenths, and the tritone, are categorized as dissonant intervals (Hutchinson and Knopoff, 1978; Huron, 2001). The classification of musical intervals in this way suggests that the ability to distinguish sensations of consonance and dissonance is important for music perception.

Consonance and dissonance have been investigated as phenomena not only of musical relevance, but also of psychoacoustic interest. In a purely psychoacoustic context, *sensory dissonance* refers to the degree to which a tone complex, presented in isolation, sounds dissonant, i.e., tense, rough, or unpleasant, while *sensory consonance* refers to a restful, smooth, or pleasant quality (Terhardt, 1984). Sensory consonance and dissonance provide the basis for the classification of intervals as *musically* consonant or dissonant. However, this relationship is not completely straightforward. For example, a minor third, which is considered a musically consonant interval, sounds dissonant when played in a very low register.

Sensory dissonance is generally regarded to be a consequence of the imperfect frequency resolution of the basilar membrane (Greenwood, 1991). According to von Helmholtz

^{a)}Portions of this work were presented at the 147th meeting of the Acoustical Society of America, New York, NY, May 2004, and at the International Conference for Music Perception and Cognition, Evanston, IL, August 2004.

^{b)}Current address: Department of Communication Sciences, University of Connecticut, Storrs, CT 06269; electronic mail: jennifer.tufts@uconn.edu

(1877/1954), sensory dissonance results from the perception of fast beats between two tones that are closely spaced in frequency. The beating tones, which are unresolved on the basilar membrane, produce amplitude fluctuations within a single auditory channel. These amplitude fluctuations create sensations of roughness and dissonance (Terhardt, 1978). Conversely, two tones that are spaced further apart in frequency do not interact within a single auditory filter to produce dissonance.

Several researchers have investigated the relationship between sensory dissonance and the frequency separation of two simultaneous pure tones in normal-hearing listeners (Plomp and Levelt, 1965; Plomp and Steeneken, 1968; Kameoka and Kuriyagawa, 1969a). In these studies, listeners evaluated the sensory consonance or dissonance of pure-tone pairs that were created either by fixing the frequency of one pure tone and varying the frequency of another, higher-frequency pure tone, or by varying the frequencies of two pure tones around a common geometric mean. These studies indicated that, across a wide frequency range, the sensory dissonance of two simultaneous pure tones is a relatively smooth function of their frequency separation in Hz. At a separation of 0 Hz, the two pure tones exactly coincide in frequency and therefore produce no dissonance. As the separation of the pure tones increases slightly from 0 Hz, beats are heard, first as fluctuations in loudness, then as a sensation of roughness. As a result, sensory dissonance increases rapidly with increasing frequency separation from 0 Hz. The frequency separation at which maximal dissonance is reached occurs at about 25%–40% of the critical bandwidth in the frequency region of the tones (Plomp and Levelt, 1965; Greenwood, 1991). As the separation between the pure tones widens further, the sensation of roughness subsides and sensory dissonance decreases. At approximately one critical bandwidth (Plomp and Levelt, 1965; Plomp and Steeneken, 1968; Greenwood, 1991), the two tones are resolved on the basilar membrane and produce a smoother sensation, with minimal dissonance. Further increases in frequency separation produce relatively little change in perceived sensory dissonance. This pattern has been observed for pure-tone pairs with lower-frequency components ranging from approximately 125 to 7000 Hz (Plomp and Levelt, 1965; Kameoka and Kuriyagawa, 1969a).

Unlike pure tone pairs, most musical sounds have many components that may interact to produce sensory dissonance. The sensory dissonance created by simultaneous harmonic complex tones is of particular interest, because the singing voice and the sounds of many musical instruments have harmonic complex spectra. While the sensory dissonance of pure-tone pairs is a relatively smooth function of their frequency separation, the dissonance of harmonic complex tone pairs varies markedly depending on their fundamental frequency (F_0) ratios and their amplitude spectra (Plomp and Levelt, 1965). Specifically, intervals with small-integer F_0 ratios (such as the octave, with an F_0 ratio of 2:1) sound consonant, while intervals with larger-integer F_0 ratios (such as the minor second, with an F_0 ratio of 16:15) sound dissonant. Small-integer ratios sound consonant ostensibly because several of the partials of the two tones are either iden-

tical in frequency, or are spaced sufficiently far apart that they do not create roughness. On the other hand, larger-integer F_0 ratios sound dissonant because many of the partials are noncoinciding and closely spaced. Given this explanation of the relationship among consonance, dissonance, and F_0 ratio, it follows that the use of consonance and dissonance in music is predicated on normal or near-normal peripheral frequency resolution. For this reason, the perception of sensory consonance and dissonance by people with SNHL is of interest from a psychoacoustic as well as a music-perception standpoint.

Deficits in frequency resolution, which often coexist with SNHL, may cause the components of a tone complex to interact over a wider frequency range than in normal-hearing listeners. Therefore, listeners with SNHL may perceive an increase in sensory dissonance between some components of a tone complex that would ordinarily sound consonant for normal-hearing listeners. This in turn may have the effect of changing the relative dissonance of musical intervals, or reducing the perceptual contrast between consonant and dissonant intervals. To the best of our knowledge, no previous studies have examined the perception of consonance and dissonance by people with SNHL.

The present study addressed two questions. First, do people with SNHL judge the sensory dissonance of musical intervals differently than do normal-hearing people? Second, are judgments of sensory dissonance by normal-hearing and hearing-impaired listeners consistent with a relationship between peripheral frequency selectivity and dissonance perception? To answer these questions, normal-hearing (NH) and hearing-impaired (HI) subjects were recruited to complete two tasks. In the first task, subjects judged the relative sensory dissonance of musical intervals created with pure tones and with harmonic complex tones. In the second task, subjects' thresholds for pure tones at 500 and 2000 Hz were measured in the presence of a notched noise. The latter task was designed to assess their peripheral frequency resolution.

II. METHOD

A. Subjects, test environment, and order of procedures

Eight subjects participated in the study. Four subjects (1 M, 3 F; mean age=50 years, range=31–63) had normal hearing in the test ear (i.e., air-conduction thresholds ≤ 20 dB HL from 0.25 to 4 kHz re: ANSI, 1996). Each of the other four subjects (2 M, 2 F; mean age=69 years, range=61–80) had a mild to moderate sensorineural hearing loss in the test ear (i.e., air-conduction thresholds between 30 and 60 dB HL at audiometric frequencies from 0.25 to 4 kHz, air-bone gaps of ≤ 10 dB from 0.5 to 4 kHz, and a normal tympanogram). Figure 1 shows the mean air conduction thresholds of the subjects. According to self-report, none of the subjects had had training in music theory, none had perfect pitch, and none was able to recognize musical intervals by ear. Thus, it was assumed that the subjects' performance was not influenced by specialized knowledge of musical intervals.

All testing took place in a double-walled sound-treated

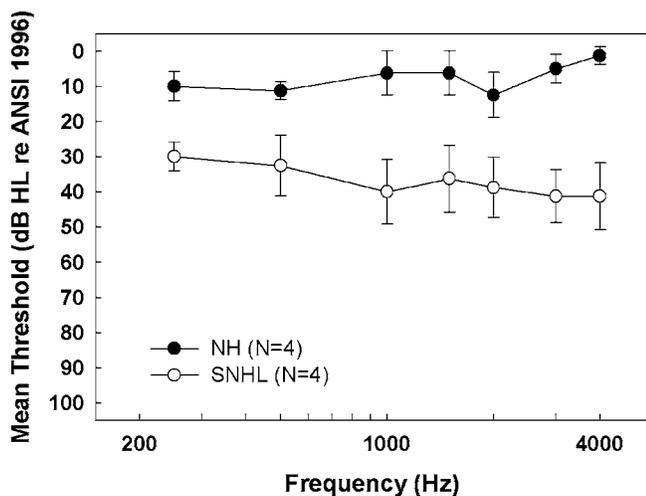


FIG. 1. Mean air-conduction thresholds (in dB HL re: ANSI 1996) for the subjects with normal hearing and the subjects with SNHL. Error bars represent standard deviations around the means.

booth. Headphone output levels were calibrated with a 500 Hz pure tone each day before data collection. Prior to enrollment in the study, each subject signed an informed consent form and a privacy statement outlining the potential uses of his or her identifying information. The subject's audiogram and tympanogram were obtained upon enrollment. Next, the subject completed the sensory dissonance judgment task, followed by the threshold-in-noise task. All stimuli were presented to the same ear. Total participation time was approximately 8 h per subject, spread over four or five sessions.

B. Sensory dissonance judgments

Subjects judged the relative sensory dissonance of pairs of tones, called *dyads*, formed by combining either two pure tones or two harmonic complex tones. The pure-tone dyads fell within two frequency regions, one centered around 500 Hz and the other centered around 2000 Hz, while the F0's of the harmonic complex dyads were centered around 500 Hz. The intervals formed by the dyads included all of the equal-tempered musical intervals of the Western tradition.

1. Stimuli

All dyads were generated digitally and played through a 16 bit digital-to-analog converter (TDT DD1) at a rate of 40 000 samples/s. The stimuli were passed through an attenuator (TDT PA4) and a headphone buffer (TDT HB6) to one channel of a set of calibrated circumaural earphones (Sennheiser, HD540). *Pure-tone* dyads (PT dyads) were composed of two simultaneous pure tones with equal amplitudes and phases drawn randomly from a uniform distribution. *Harmonic complex* dyads (HC dyads) were composed of two simultaneous harmonic complex tones, each having six components (F0 and five harmonics). All twelve components were of equal amplitude and random phase. Each dyad was 750 ms in total duration, including 50 ms raised-cosine onset and offset ramps.

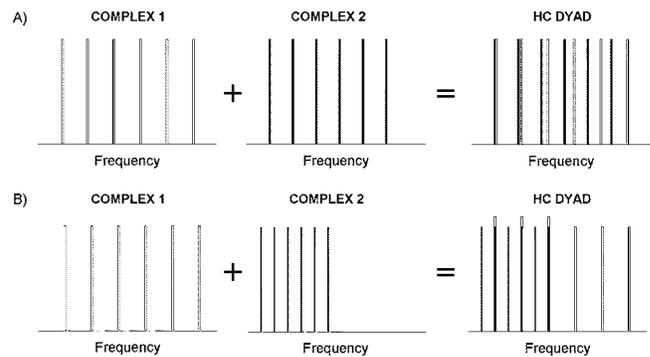


FIG. 2. Schematic example of the construction of two harmonic complex (HC) dyads. (A) HC dyad composed of two harmonic complex tones with a fundamental frequency ratio of 16:15 (a minor second). (B) HC dyad composed of two harmonic complex tones with a fundamental frequency ratio of 2:1 (an octave).

The overall presentation level of each dyad was approximately 83 dB SPL.¹ This level was chosen so that each dyad component would have a sensation level of at least 10 dB for each HI listener, without creating discomfort for either the NH or HI listeners. The levels of the individual components of the dyads were 80 and 72 dB SPL for the PT and HC dyads, respectively.

Two sets of 26 PT dyads were created. The dyads within a set were centered at a geometric mean frequency of either 500 Hz (PT 500 Hz dyads) or 2000 Hz (PT 2000 Hz dyads). In each set, the frequency separation between the dyad components was varied in quartertone steps (i.e., equal logarithmic steps of $2^{1/24}$), beginning at 0 Hz. With this step size, the thirteen equal-tempered musical intervals from the unison to the octave were represented in each set, as well as the thirteen interposed intervals ranging from the quartertone above the unison to the quartertone above the octave. One set of 26 HC dyads was also created (HC 500 Hz dyads). The F0's of the HC 500 Hz dyad components were identical to the frequencies of the PT 500 Hz dyad components. Figure 2 depicts a schematic example of the construction of two HC dyads whose F0's form a ratio of either (A) 16:15 (a minor second) or (B) 2:1 (an octave). Notice that the resulting dyads may have fewer than 12 components if the two added complexes have some partials in common.

Table I lists the PT dyads' musical interval names (where applicable), frequency ratios, and component frequencies. It should be noted that all of the equal-tempered musical intervals except the unison and the octave deviate slightly from simple integer ratios. For example, the equal-tempered fifth deviates from a ratio of 3:2 by two cents (1/600 of an octave). However, such small deviations do not reduce the acceptability of these intervals in musical practice (Vos, 1988).

2. Procedure

Subjects judged the relative sensory dissonance of the dyads within each set via a paired comparison task. On each trial, the subject heard two dyads from the same set (either the PT 500 Hz, PT 2000 Hz, or HC 500 Hz dyads), separated by a 500 ms silent interval, and chose the dyad that sounded more unpleasant. The term "unpleasant" was de-

TABLE I. The pure-tone dyads are shown in order of increasing interval width. Included are the musical interval names of the dyads (where applicable), their equal-tempered frequency ratios, and their component frequencies. The just integer ratios of the musical intervals are shown in parentheses. The fundamental frequencies of the components of the harmonic complex dyads are identical to the frequencies of the pure-tone dyads centered at 500 Hz.

Dyad No.	Musical interval name	Frequency ratio	Pure-tone frequencies in Hz for dyads centered at 500 Hz	Pure-tone frequencies in Hz for dyads centered at 2000 Hz
1	Unison	1.000 (1:1)	500–500	2000–2000
2		1.029	493–507	1971–2029
3	Minor second	1.059 (~16:15)	486–515	1943–2059
4		1.091	479–522	1915–2089
5	Major second	1.122 (~9:8)	472–530	1888–2119
6		1.155	465–537	1861–2150
7	Minor third	1.189 (~6:5)	459–545	1834–2181
8		1.224	452–553	1808–2213
9	Major third	1.260 (~5:4)	446–561	1782–2245
10		1.297	439–569	1756–2278
11	Perfect fourth	1.335 (~4:3)	433–578	1731–2311
12		1.374	427–586	1706–2344
13	Tritone	1.414 (~45:32)	420–595	1682–2378
14		1.456	414–603	1658–2413
15	Perfect fifth	1.498 (~3:2)	409–612	1634–2448
16		1.542	403–621	1611–2484
17	Minor sixth	1.587 (~8:5)	397–630	1587–2520
18		1.634	391–639	1565–2557
19	Major sixth	1.682 (~5:3)	386–648	1542–2594
20		1.731	380–658	1520–2631
21	Minor seventh	1.782 (~9:5)	375–667	1498–2670
22		1.834	369–677	1477–2709
23	Major seventh	1.888 (~15:8)	364–687	1458–2748
24		1.943	359–697	1435–2788
25	Octave	2.000 (2:1)	354–707	1414–2828
26		2.059	349–717	1394–2870

scribed as synonymous with “rough, dissonant, or discordant,” as opposed to “pleasant, smooth, pure, or harmonious.” Each dyad was paired twice with every other dyad in its set, once in each order, for a total of 650 trials. For each listener, the order of these 650 trials was randomized and then split to create two blocks of 325 trials each. This was done for each dyad set (PT 500 Hz, PT 2000 Hz, and HC 500 Hz dyads), giving a total of six blocks. The order of these six blocks was randomized for each listener.

A running score was kept for each dyad during testing. Initially, each dyad was assigned a score of zero. Each time a particular dyad was chosen as more unpleasant in a trial, 0.5 was subtracted from its score. For example, if a particular dyad was always chosen as more unpleasant in all of its pairings with the other dyads in its set, it would obtain a score of $-0.5 * 50 = -25$, the lowest possible score. If a particular dyad was never chosen as more unpleasant in any of its pairings, its score would remain at zero, the highest possible score. At the completion of all six blocks, the subject had three sets of scores, each one representing the relative dissonance of the dyads within a stimulus set.

C. Estimation of peripheral frequency selectivity

Auditory filter bandwidths for each subject were estimated for signal frequencies of 500 and 2000 Hz using a

notched-noise masking task (see Patterson and Moore, 1986, for a discussion of this method of estimating frequency resolution). This procedure involved obtaining several masked threshold estimates, with masking provided by two bands of noise, one above and one below the signal in frequency. As the frequency separation between the signal and the masking bands increased, the noise level required to mask the constant-level signal increased. The rate of increase of the masker level was used to estimate the shape of the auditory filter centered at the signal frequency.

1. Stimuli

The signals were tones of either 500 or 2000 Hz, selected to correspond to the geometric means of the dyads in the sensory dissonance judgment task. Signal duration was 300 ms, with 50 ms cosine-squared onsets and offsets. Notched-noise maskers consisted of two bands of noise positioned on either side of the signal frequency. The bandwidths of the noises were 0.4 times the signal frequency, with steep skirts (falling more than 75 dB per 100 Hz). The duration of the noise maskers was 400 ms, with the signal temporally centered in the noise. Eight notched-noise maskers were generated for each signal frequency, six with the notch centered on the signal frequency (i.e., symmetric

notches) and two with either the upper or lower noise band placed closer to the signal frequency (asymmetric notches).

Signals and notched-noise maskers were generated digitally, and played out at 40 000 samples/s through separate channels of a digital-to-analog converter (TDT DD1). The signal and noise channels were passed through separate programmable attenuators (TDT PA4), and added together (TDT SM3) before being passed through a headphone buffer (TDT HB6) to one channel of a set of TDH-49P earphones. Signal levels were fixed at 60 dB SPL for the NH listeners, and either 70 or 80 dB SPL for the HI listeners. Noise level was varied adaptively to determine the level that just masked the signal.

2. Procedure

Each masked threshold was measured using a modification of the single-interval yes-no maximum-likelihood adaptive procedure described by Green (1993). This procedure has been shown to produce reliable threshold estimates in 12 to 25 trials, depending on the number of catch trials included in the procedure (Gu and Green, 1994; Leek *et al.*, 2000). On each trial, the signal and noise were presented together, with the signal level fixed across trials and the notched-noise level determined by an adaptive track. Each block of trials began with the noise level low enough so that the signal was clearly heard. After each presentation, the subject indicated by a button press on a response box whether the signal was heard or not heard. A light on the box indicated that the response had been registered, but no other feedback was given. Catch trials, in which no signal was present, occurred on 20% of the presentations, and the responses to those trials were used to estimate a false alarm rate. After each presentation and response, a set of candidate psychometric functions was consulted. All previous responses in the block were used to determine which of the candidate functions was the most likely to represent the data collected up to that point. The selected function was then entered at 70% to determine the masker level for the next trial. Presentations continued until the confidence interval for the current threshold value was less than 1 dB. At the end of the threshold track, the level of noise necessary to produce 70% correct detections of the signal was extracted from the final estimated psychometric function. The average of three such measurements was taken as the threshold for that block of trials, and the average of two blocks for each notched-noise condition constituted the final threshold value. In cases where thresholds from the two blocks differed by more than 3 dB, an additional block was run and all threshold estimates were averaged.

While this adaptive procedure is not unbiased like a forced choice procedure, the implementation here has a number of characteristics that increase its reliability and validity. The starting level for each adaptive run was selected randomly over a range of values that produced a clearly audible signal; each track continued until a criterion variability measure was reached; and thresholds were estimated from a minimum of six adaptive thresholds (two blocks of three tracks each), with additional threshold runs if the two block estimates differed by more than 3 dB. The reliability of this procedure has been discussed extensively by Green and his

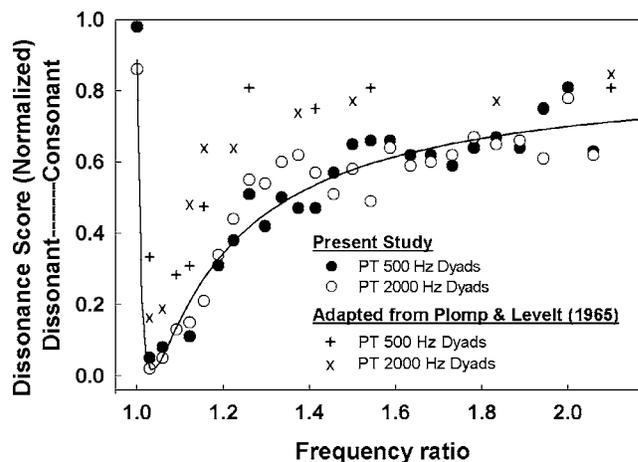


FIG. 3. Median normalized sensory dissonance scores of the NH group ($N=4$) for each of the two pure-tone dyad sets, plotted as a function of the frequency ratio of the dyad components. Closed circles represent the data for the pure-tone dyads centered at 500 Hz; open circles represent the data for the pure-tone dyads centered at 2000 Hz. The two sets of data are fitted by a single lognormal curve (see the text). Selected data from Plomp and Levelt (1965), translated to these axes, are shown for comparison. Crosses represent data for the pure-tone dyads centered at 500 Hz; Xs represent data for the pure-tone dyads centered at 2000 Hz.

colleagues and others (e.g., Gu and Green 1994; He *et al.*, 1998; Florentine *et al.*, 2000; Leek *et al.*, 2000).

III. RESULTS

A. Sensory dissonance judgments

Each subject produced three sets of dissonance scores, each one representing the relative sensory dissonance of the dyads within a stimulus set (PT 500 Hz, PT 2000 Hz, or HC 500 Hz dyads). Within each subject group (NH versus HI), the median dissonance score of each dyad, and its associated interquartile range, were calculated. Scores could range from 0, indicating maximum consonance, to -25 , indicating maximum dissonance. For the NH subjects, the average interquartile ranges were 1.6 for the PT 500 Hz dyads, 4.7 for the PT 2000 Hz dyads, and 3.4 for the HC 500 Hz dyads. HI subjects had average interquartile ranges of 4.4, 2.3, and 3.8 for the corresponding dyad sets.

The median scores for each dyad were normalized between 0 and 1, with 1 representing maximal sensory consonance and 0 representing maximal sensory dissonance. The normalized scores of the PT 500 Hz and PT 2000 Hz dyads were plotted as a function of the frequency ratio of the dyad components; the normalized scores of the HC 500 Hz dyads were plotted as a function of the F_0 ratio of the dyad components.

Figure 3 shows the scores of the PT 500 Hz and PT 2000 Hz dyads for the NH subjects. Plotted as a function of frequency ratio, the two sets of scores are nearly superimposed. The pattern of the data is similar to that observed in previous studies of the sensory dissonance of PT dyads. Specifically, as the frequency separation of the dyad components widens from 0 Hz, dissonance rapidly increases to a maximum, then decreases somewhat more gradually, eventually reaching a plateau. Selected points from Plomp and Levelt

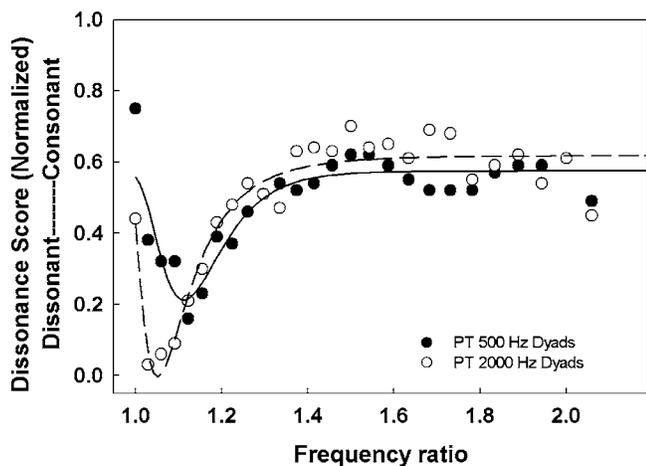


FIG. 4. Median normalized sensory dissonance scores of the HI group ($N=4$) for each of the two pure-tone dyad sets, plotted as a function of the frequency ratio of the dyad components. Closed circles represent the data for the pure-tone dyads centered at 500 Hz; open circles represent the data for the pure-tone dyads centered at 2000 Hz. Each of these two sets of data is fitted by a lognormal curve (see the text).

(1965) are included in Fig. 3 for comparison. These points represent approximate translations of Plomp and Levelt's (1965) data, which were obtained with a different methodology and different instrumentation than were used in the present study. Nevertheless, they show agreement with the current data in their overall pattern.

The two sets of dissonance scores of the NH group were initially fit by separate lognormal functions.² Comparison of these functions using Akaike's Information Criterion (AIC; Akaike, 1974) indicated that the two sets of scores could be represented by a single function fit to the combined data ($R^2=0.93$; corrected AIC values for one versus two functions, respectively: 58.25 and 65.93). The AIC compares maximum likelihood estimates of competing functions adjusted for the number of free parameters. The function associated with the smaller AIC value provides the better fit to the data. This function is shown as a solid line in Fig. 3. Additionally, lognormal functions were fit to the pure-tone dissonance scores of each individual NH subject, and are used in the analysis of the relationship of auditory filter bandwidth to sensory dissonance perception, described in Sec. III B.

Figure 4 shows the normalized scores of the PT 500 Hz and PT 2000 Hz dyads for the HI subjects. As is the case for the NH group, maximal dissonance occurs at a narrow frequency separation, followed by a decrease in dissonance to a plateau as the frequency separation increases. The two sets of dissonance scores of the HI group were fit by separate lognormal functions, shown as solid (PT 500 Hz dyads) and dashed (PT 2000 Hz dyads) lines in Fig. 4. Statistical evaluation using the AIC indicated that the two sets of scores were better fit by these two functions rather than one function for the combined data ($R^2=0.81$ and 0.91 for the PT 500 Hz and PT 2000 Hz curves, respectively; corrected AIC values for one versus two functions, respectively: 86.69 and 65.62). As was done for the NH subjects, lognormal functions were also fit to the pure-tone dissonance scores of each individual HI

subject, with the exception of one subject's PT 500 Hz scores, which could not be adequately represented by a lognormal function.

Visual inspection of the pure-tone dissonance curves shown in Fig. 3 (NH) and Fig. 4 (HI) suggest several differences between the listener groups. First, the overall range of median dissonance scores for the HI listeners is compressed, relative to the range of scores for the NH listeners. While the normalized dissonance scores of the NH group span nearly the entire possible range, from 0.02 to 0.98, the normalized scores of the HI group span a smaller range from 0.03 to 0.75.

One aspect of the reduced range may be observed in the scores for the unison (the dyad with a 0 Hz frequency separation). The HI listeners did not judge the unison to be as consonant as did the NH listeners. For the NH group, the unison has median scores of 0.98 and 0.86, respectively, for the PT 500 Hz and PT 2000 Hz dyad sets, versus 0.75 and 0.44, respectively, for the HI group. A two-way analysis of variance with one repeated factor (frequency region) was carried out on the log-transformed individual unison scores across the two groups of listeners. The difference between the groups in the dissonance scores at the unison was statistically significant [$F(1,6)=7.20, p<0.04$]. In addition, the main effect of frequency region was statistically significant [$F(1,6)=15.89, p<0.01$]. The interaction between listener group and frequency region was not significant ($p>0.20$). These analyses indicate that, in general, the 500 Hz unison was more consonant than the 2000 Hz unison for both groups, but the unison at each frequency was perceived as less consonant by the HI listeners compared with the NH listeners.

Finally, maximal sensory dissonance (i.e., the minima of the dissonance curves) occurs at a larger frequency ratio on the PT 500 Hz curve of the HI listeners than on the other three pure-tone dissonance curves. Specifically, on the PT 500 Hz curve of the HI group, the point of maximal dissonance falls near the major second; on the PT 2000 Hz curve of the HI group, and both pure-tone dissonance curves of the NH group, the points of maximal dissonance occur at intervals smaller than the minor second. The frequency ratios at which maximal sensory dissonance occurred for each individual and frequency region were entered into a two-way analysis of variance, with frequency region as a repeated factor. The differences noted above were confirmed statistically: there was no main effect of group [$F(1,6)=3.008, p>0.10$], but there was a significant main effect of frequency region [$F(1,6)=13.293, p=0.01$]. This main effect was primarily due to the significant interaction between group and frequency [$F(1,6)=7.78, p=0.03$].

The closed symbols in panels (A) and (B) of Fig. 5 show the median scores of the HC 500 Hz dyads of the NH and HI listeners, respectively (open symbols will be discussed later). As the separation between the F_0 's of the dyad components widens from 0 Hz, the dissonance of the HC dyads rapidly increases to a maximum at an interval of one quartertone and then decreases somewhat more gradually for both groups of listeners. However, unlike the pure-tone sensory dissonance curves, the harmonic complex sensory dissonance curves ex-

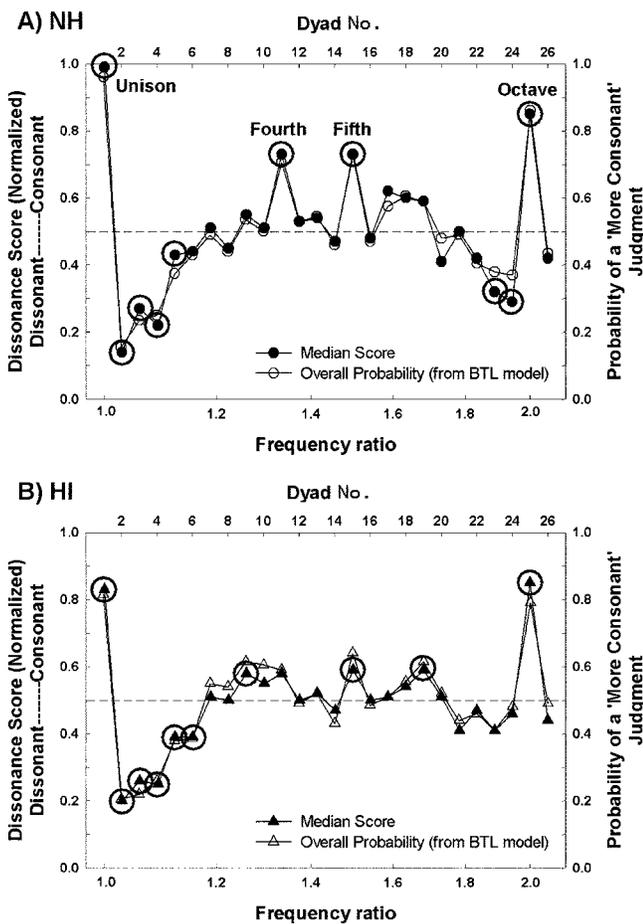


FIG. 5. Median normalized sensory dissonance scores of the NH (A) and HI (B) groups for the harmonic complex dyad set, shown as closed symbols and plotted as a function of the fundamental frequency ratio of the dyad components. Those scores that are circled are significantly more or less dissonant than predicted by chance ($p < 0.05$), which is signified by the horizontal dashed line on each panel. [Note that the tests of difference from chance were conducted on pooled raw scores rather than the median normalized scores depicted here.] The open symbols indicate the overall probability of each dyad being judged more consonant than the other dyads in the set, based on the BTL model of paired comparisons. These data should be referred to the right axis.

hibit several sharp peaks associated with small-integer FO ratios. This observation is consistent with previous research on the dissonance of HC dyads (e.g., Kameoka and Kuriyagawa, 1969b).

These data will be evaluated in two ways. First, one may ask which of these dyads were perceived as clearly dissonant or clearly consonant. This question will be addressed using a chi-squared analysis of differences of the scores from chance performance. Next, an analysis will be provided that takes into account the influence of all of the dyads in a set on individual judgments in determining the relative consonance or dissonance of the dyads.

Tests for paired comparison data (David, 1988) were used to determine which dyads were more consonant or dissonant than would be expected based on chance performance. The null hypothesis of no difference from chance was evaluated for each dyad using a normal approximation to the chi-square distribution. For each group of four listeners (with two replications each), the pooled scores were

tested against a criterion value calculated for a significance level of 0.05, corrected for the 26 different significance tests to be performed, and divided by two for a two-tailed test of difference from chance (represented on the panels of Fig. 5 as a horizontal line at a score of 0.5). The dyads whose scores were significantly different from chance are circled in Fig. 5. These dyads were perceived as either highly dissonant or highly consonant, relative to the more ambiguous perceptions that are reflected by scores near chance.

As shown in Fig. 5, the NH group identified the intervals of the unison, octave, fifth, and fourth as highly consonant. Further, they judged the intervals equivalent to or smaller than a major second (dyads 2–5), and intervals near the octave (dyads 23 and 24), to be significantly dissonant. The HI group identified as highly consonant the unison, the octave, and the fifth, as well as the major third and major sixth, but not the fourth. HI listeners judged intervals smaller than a minor third (dyads 2–6) to be significantly dissonant, but did not judge dyads near the octave to be significantly dissonant.

Analyzing differences from chance performance does not directly take into account the influence of all the other dyads on individual comparisons between two dyads. For a comparison between a given pair of dyads, the probability of a particular outcome will be influenced by the other dyads in the stimulus set. A model of paired comparisons developed by Bradley and Terry (1952), further modified by Luce, and described in David (1988), was implemented to generate these probabilities. A similar analysis was conducted by Pressnitzer and McAdams (1999) to study the perception of roughness, as well as by Uppenkamp *et al.* (2001) to evaluate paired comparisons of “compactness” (an aspect of timbre).

The Bradley-Terry-Luce (BTL) model is usually described with reference to rankings of teams in a baseball league (see Agresti, 1990, for an intuitive description of this model of paired comparisons). The number of wins is not only dependent on a team’s own ability, but also on the abilities of each of the other teams (see David, 1988). The BTL model takes these dependencies into account to establish the probability of a win for each team when playing every other team, as well as an overall probability of winning. This model was used here to establish the probabilities of the possible outcomes of all individual head-to-head comparisons between the dyads. These individual probabilities were combined so that an overall probability of “winning” (i.e., being judged more consonant in relation to all other dyads) was generated for each dyad. In effect, the analysis converted the individual judgments into a ranking of all of the dyads along the consonance/dissonance dimension. These overall probabilities were normalized to sum to 1.0. They are shown as the open symbols in Fig. 5, referred to the right-hand ordinate. These constructed probabilities, which take into account the scores of all of the dyads, closely follow the pattern of median scores.

For each dyad, a probability function was constructed that displays its probability of “winning” when paired with every other dyad. The 26 probability functions are shown in Fig. 6(A) (NH) and Fig. 6(B) (HI). Each curve on these panels represents the probability (shown along the ordinate)

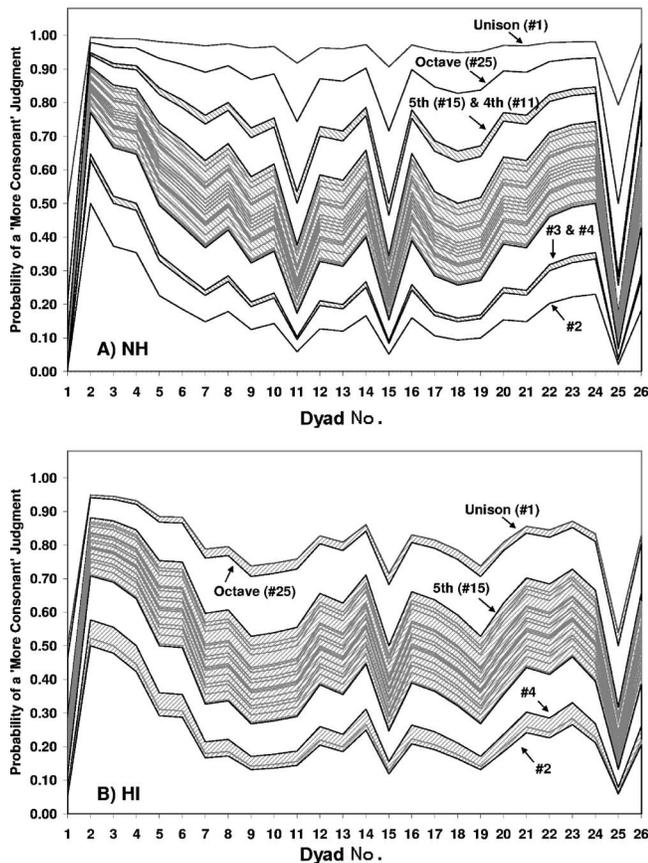


FIG. 6. Probability functions for each dyad, based on the BTL model of paired comparisons. Each curve shows the individual probabilities of that dyad being judged more consonant in relation to each of the other 25 dyads, shown on the abscissa. Curves whose average probabilities are less than one standard deviation different from neighboring curves are grouped together by shading (see the text for a more complete explanation). (A) and (B) Data from NH and HI listeners, respectively. There are six groupings of the probability functions in (A) and three groupings in (B) indicating that NH listeners more clearly separated the dyads in terms of their relative consonance/dissonance than did the HI listeners.

that the given dyad was perceived as more consonant when compared against every other dyad (shown along the abscissa), according to the BTL model. For example, in Fig. 6(A), the probability of the perfect fifth (dyad 15) being chosen as more consonant is 0.80 in its pairing with dyad 8, and 0.28 in its pairing with the octave (dyad 25).

The curves are arranged in Fig. 6 in order from high to low average probability, where the average probability for each dyad was computed across all of the values along its curve. It is apparent from Fig. 6 that the curves are not equally distributed along the probability axis. Instead, some appear to group together, reflecting those dyads that share similar consonance patterns. Groupings were defined on an ad hoc basis by taking the mean of all of the differences in average probabilities between adjacent curves, and grouping together those curves whose average probabilities did not differ from neighboring curves by more than one standard deviation of the mean. Groupings are demarcated in the figure by shading.

There are six distinct groupings for the NH listeners and three groupings for the HI listeners. For the NH listeners [Fig. 6(A)], the unison and the octave are not grouped with

any other intervals or with each other. The fifth and fourth are grouped together. These dyads are very similar to one another in having lower average probabilities than the octave, but higher average probabilities than the rest of the dyads. The large shaded area in the middle of Fig. 6(A) groups together all but three of the remaining dyads. Two of these, the minor second and its upper neighboring dyad (dyads 3 and 4), are grouped together; the quartertone interval (dyad 2) has the lowest average probability and is separate from the rest, indicating that it is least consonant overall.

As shown in Fig. 6(A), the probability functions of the unison and the octave are fairly flat in comparison with the probability functions of the other dyads. The consistently high probabilities of the unison and octave indicate that these two dyads are very resistant to being chosen as more dissonant, no matter which dyad they are paired with. The only intervals that appear to affect the unison and octave, aside from each other, are the fifth and fourth. This is shown by the dips in the probability functions at the fifth and fourth (dyads 15 and 11, respectively, on the abscissa).

Compared with the NH group, the HI subjects produced fewer distinct groups of dyads. As shown in Fig. 6(B), the highly consonant unison and octave form a group, as do the highly dissonant minor second and its two neighbors (dyads 2–4). All of the 21 remaining dyads are grouped together. The probability functions of the unison and octave are not as flat as those of the NH group. Dips in the functions occur at the fifth (dyad No. 15) [and the octave (dyad No. 25) and unison (dyad No. 1), respectively], but not the fourth (dyad No. 11). The HI listeners apparently perceived the fifth as a relatively consonant interval, although neither it nor the fourth were grouped separately from the other dyads.

Overall, these results suggest that the HI listeners do not distinguish intervals in terms of consonance and dissonance as clearly as do the NH listeners. This is consistent with the more compressed puretone dissonance curves of the HI listeners, noted earlier.

B. Relationship of dissonance to auditory filter bandwidths

The thresholds generated by the notched-noise procedure for each individual were used to estimate auditory filter shapes. The auditory filters were derived using the polyfit procedure described by Rosen and Baker (1994). A least-squares fitting procedure was implemented to find the filter weighting function that best predicted the set of thresholds, given the assumptions of the power spectrum model of masking. The general weighting function has the form

$$W(g) = (1 - pg)e^{-pg}, \quad (1)$$

where g is a normalized frequency variable, representing the difference between the center frequency and a given point on the filter skirt, and p determines the passband and the slope of the filter skirt (Patterson and Moore, 1986). The skirt parameter, p , was fit separately on either side of the filter. Equivalent rectangular bandwidths (ERBs) of the individual auditory filters, given as a proportion of the center frequency, were computed as $[2/p(\text{lower skirt})]$

$+[2/p(\text{upper skirt})]$, as described by Glasberg and Moore (1990).

The mean relative ERBs of the auditory filters at 500 Hz were 0.20 and 0.26 for the NH and HI listeners, respectively. At the center frequency of 2000 Hz, the mean relative ERBs were 0.21 and 0.29 for the two groups, respectively. A *t*-test at each center frequency indicated that the difference in ERB between groups was statistically significant at 500 Hz ($p = 0.002$), but not at 2000 Hz ($p > 0.3$), where there was considerable variability among the HI listeners. Large variability in estimates of auditory filter bandwidths of HI listeners has been noted frequently before (e.g., Glasberg and Moore, 1986; Peters and Moore, 1992). Other factors that could account for the similar auditory filter bandwidths for NH and HI listeners at 2000 Hz are the high levels at which the measurements were made, and the relatively mild hearing losses of most of the HI subjects. Auditory filters of NH listeners measured at high stimulus levels are typically broader than those measured at lower levels (Leek and Summers, 1993). Given listeners with normal hearing and a signal level of 60 dB SPL, Rosen and Baker (1994) calculated an ERB of 0.22 at 2000 Hz, nearly identical to the mean ERB measured here. For HI listeners, Baker and Rosen (2002) reported an average ERB at 2000 Hz of about 0.25, measured at similar signal levels to those used here (the value estimated from their Fig. 5).

The significant difference in auditory bandwidths between the two groups of listeners at 500 Hz is noteworthy, given the differences observed in the dissonance curves of the PT 500 Hz dyads. Recall that previous investigators (Plomp and Levelt, 1965; Greenwood, 1991) estimated that maximum sensory dissonance occurs at frequency separations of approximately 25%–40% of the critical bandwidth, and that two tones separated by approximately one critical bandwidth are consonant. Figure 7 illustrates the strength of these two relationships in the data reported here.

Two points were extracted from each of the individual subjects' pure-tone dissonance curves for comparison with auditory filter bandwidth values. The curves fitted to the dissonance data have an exponential shape as the frequency difference between dyad components becomes large. Representing the exponential growth of the approach to more consonant responses with increasing frequency separation, a "growth constant" was extracted from each individual dissonance curve. The growth constant was defined as the abscissa value at which the curve approached an asymptotic value on the ordinate, i.e., at approximately two-thirds of the difference between the minimum and maximum ordinate values. The growth constant was used here as an estimate of the point beyond which increases in frequency separation did not result in appreciable increases in consonance. This point may be thought of as analogous to the "shoulder" in Plomp and Levelt's (1965) data. Figure 7(A) plots these growth constants as a function of the ERB in Hz for each subject at the two center frequencies. (Note that only three hearing impaired data points are shown for 500 Hz; this is because one subject's data could not be adequately fit by the lognormal function, and therefore the growth constant could not be adequately estimated for that condition and subject.) If the re-

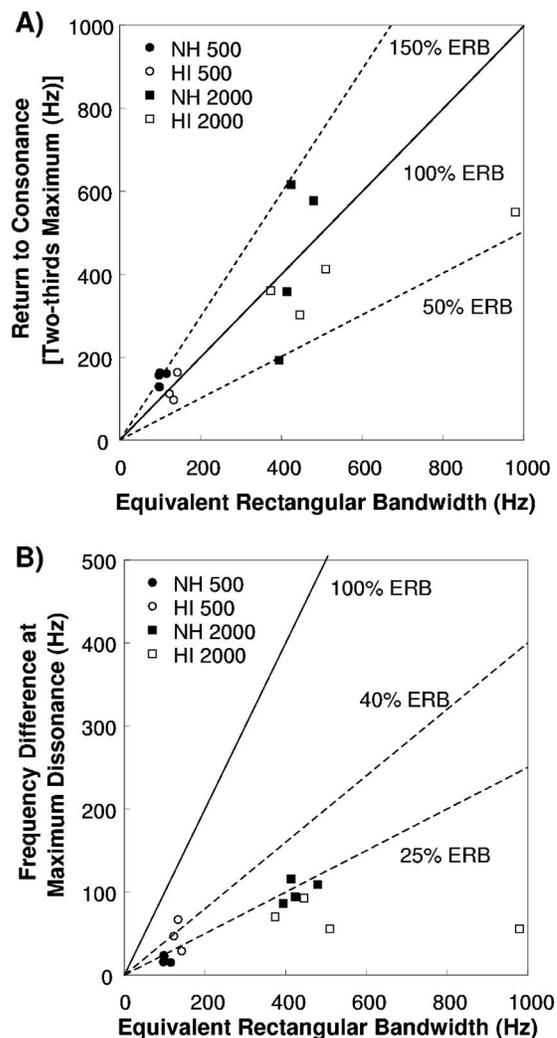


FIG. 7. (A) A measure of the return to consonance of the pure-tone dissonance curves as a function of the ERB of the auditory filters for individual subjects. NH data are shown with closed symbols; HI data are shown with open symbols. The solid line is included for reference, and represents frequency separations corresponding to 100% of the ERB. (B) Frequency separation at maximum sensory dissonance as a function of ERB for each individual subject. The solid and dashed lines are included for reference, and represent frequency separations corresponding to 100%, 40%, and 25% of the ERB, respectively.

turn to consonance occurs at approximately one critical bandwidth, and if the growth constant and ERB are reasonable approximations of these two phenomena, then the data points should fall on the main diagonal, labeled 100% ERB, in Fig. 7(A). For the NH listeners, the growth constant occurred at an average of 149% of the ERB at 500 Hz, and 100% of the ERB at 2000 Hz. For the HI listeners, the percentages were 93% and 75%, respectively. Thus, to a first approximation, the relationship between critical bandwidth and the return to consonance is observed here.

The other noted relationship between auditory filter bandwidth and sensory dissonance perception is that maximum dissonance will occur when the two components of a dyad are separated by 25%–40% of a critical band. For the current data, this relationship was evaluated by first finding the minimum of each of the individual fitted curves for the PT 500 Hz and PT 2000 Hz dyad sets. Next, as shown in

Fig. 7(B), the frequency separation corresponding to the minimum of each curve was plotted as a function of the individual subject's measured ERB at the appropriate center frequency. The solid line on the panel, with a slope of one, represents the hypothetical case that maximal dissonance occurs exactly at the ERB value. The dashed lines on the panel indicate 25% and 40% of the measured ERBs. As expected, all the data fall below the 100% ERB line, indicating that maximum sensory dissonance is perceived at frequency separations that fall well within a single auditory filter for both NH and HI listeners. Most data are near or below the 25% ERB line. For the NH subjects, maximum dissonance occurred at an average of 19% of the ERB at 500 Hz, and 24% of the ERB at 2000 Hz. For the HI subjects, maximum dissonance occurred at an average of 36% of the ERB at 500 Hz, and 14% of the ERB at 2000 Hz. Interestingly, for NH listeners, the point of maximal dissonance is a more nearly constant proportion of the ERB across frequency regions than it is for the HI listeners.

IV. DISCUSSION

A. Sensory dissonance of pure-tone dyads

The pure-tone dissonance curves measured in this study were similar to one another in their general characteristics: a relatively consonant score for the unison (with the exception of the PT 2000 Hz curve of the HI group), followed by a dip to maximum dissonance and a subsequent return to consonance as the frequency difference between dyad components increased. Furthermore, the predicted relationships between critical bandwidth on the one hand, and maximum dissonance and the return to consonance on the other hand, were roughly upheld by the data (although the small number of subjects and the subjective nature of the dissonance judgment task probably contributed to the imprecise nature of the relationship seen here).

Some important differences were evident in the pure-tone dissonance curves of the NH and HI groups. First, the HI group did not judge the unison to be as consonant as did the NH group. In theory, the unison has zero dissonance, since it comprises two pure tones of identical frequency. Therefore, it was expected that the unison would be judged to be the most consonant interval in each stimulus set. In fact, this is what occurred for the NH listeners; however, the results of the HI listeners did not show the expected pattern. Although the HI listeners judged the unison to be the most consonant interval in the PT 500 Hz stimulus set, its score was lower than that given to the unison by the NH group. Even less expected was the finding that HI listeners judged the unison in the PT 2000 Hz stimulus set to be the seventh most *dissonant* interval in that set. It is possible that the 2000 Hz unison sounded sharper, shriller, or less "pure" for the HI subjects than for the NH listeners. Moore (2001) reported preliminary data in which subjects with hearing loss were asked to rate pure tones for their "distinctness" or "noisiness." The subjects produced scores indicating that tones in regions of hearing loss, particularly in the higher frequencies, were not as distinct as tones in regions of more normal hearing. This preliminary report provides some sup-

port for the notion that the 2000 Hz unison may have sounded slightly noisy or distorted to the HI listeners, which may account for its unexpectedly low consonance. Other evidence of distorted pitch perception in the presence of hearing loss has been reported by Larkin (1983) and by Leek and Summers (2001).

Maximal sensory dissonance fell at a larger interval on the PT 500 Hz dissonance curve of the HI group than it did on their PT 2000 Hz curve and on the two pure-tone dissonance curves of the NH listeners. Specifically, the major second (dyad 5) was maximally dissonant for the PT 500 Hz stimuli of the HI group, whereas the quartertone (dyad 2) was maximally dissonant for the PT 2000 Hz stimuli of the HI group and for both stimulus sets of the NH group (see Figs. 3 and 4). This finding is interesting in light of the two mechanisms of roughness perception proposed by Zwicker and Fastl (1990). [Recall that roughness and dissonance, while not synonymous, are closely related (Terhardt, 1974)]. Zwicker and Fastl (1990) provided data showing that maximal roughness of amplitude-modulated (AM) pure tones is frequency-dependent below about 1000 Hz. They argued, therefore, that at low frequencies, frequency selectivity is the limiting factor in roughness perception. Zwicker and Fastl (1990) also showed that for tones above about 1000 Hz, maximal roughness occurred at AM rates of approximately 75 Hz, independent of stimulus frequency. This finding suggests that, for high-frequency stimuli, roughness is not linked so strongly to critical bandwidth, but is instead limited by the ability of the ear to follow fast amplitude modulations. Given these data, it is reasonable to propose that any effects of broader auditory filters in the present study would be seen more clearly in the 500 Hz data than in the 2000 Hz data. Indeed, the significant broadening of the auditory filters at 500 Hz for the HI listeners relative to the filters of the NH listeners may be related to the shift of the point of maximal dissonance for the 500 Hz dyad set of the HI group.

B. Sensory dissonance of harmonic complex dyads

Consistent with previous research on the dissonance of HC dyads, the harmonic complex dissonance curves exhibit several sharp peaks associated with small-integer F0 ratios. For the NH group, peaks occur at the "perfect" consonances, i.e., the unison, octave, fifth, and fourth. These intervals, having coinciding or very nearly coinciding partials, sound more consonant than the immediately neighboring intervals. Another distinguishing characteristic of a perfect interval is its tendency to be perceived as highly fused. Terhardt (1974, 1984) explained the affinity of tones forming an octave, fifth, or fourth on the basis of our repeated exposure to these intervals in the spectra of harmonic complex sounds, including, especially, voiced speech. The fact that musically naïve listeners judged these intervals to be the most consonant of all the intervals in the set indicates that their special status in music is due to fundamental perceptual qualities of these intervals and is not merely a learned convention of music theory.

The harmonic complex dissonance curve of the NH group showed regions of marked sensory dissonance near the

unison, the octave, and the fifth. The existence of these regions of dissonance underscores the importance placed on preserving octaves and fifths in various tunings and temperaments (e.g., Pythagorean tuning and equal temperament, among others).

Like the NH listeners, HI listeners judged the unison and octave to be very consonant. However, the peaks at the fifth, and especially at the fourth, were not as robust. In addition, the harmonic complex dissonance curve of the HI group had a region of marked dissonance near the unison only. Together with the analysis presented in Fig. 6, these findings provide evidence that HI listeners do not distinguish the relative sensory dissonance and consonance of intervals as clearly as do NH listeners. This loss of contrast among the intervals suggests that HI listeners would not fully experience the variations in musical tension supplied by dissonant and consonant intervals.

The loss of contrast may be related to a reduction in pitch strength, which often accompanies SNHL (Leek and Summers, 2001). A reduction in pitch strength may lessen the degree of fusion of highly consonant intervals to the point that they do not contrast clearly with the more dissonant neighboring intervals. The loss of contrast may also be related to poorer frequency selectivity. If the auditory filter bandwidths of the HI listeners were generally somewhat broader across the frequency range in which the HC dyads' components fell (approximately 350–4300 Hz), as is suggested in the current data especially for the lower frequency regions, then more extensive interactions may have occurred among the components, thereby blurring distinctions in dissonance among the intervals.

It is likely that sensitivity to amplitude modulation (AM) per se does not account for the differences in the NH and HI groups seen here. Bacon and Gleitman (1992) reported that the ability of their HI subjects with relatively flat SNHL to detect AM did not differ from NH subjects when audibility was accounted for. Further, the stimuli in the present study were maximally amplitude-modulated, and presumably this modulation was detectable by both groups of listeners.

C. Effects of level

An additional factor that may affect sensory dissonance perception is the level above threshold at which the dyad components are presented. If lowering the sensation level (SL) of a dyad lessens its perceived dissonance, then this might explain the loss of contrast seen in the harmonic complex dissonance curves of the HI group. To investigate this possibility, two of the four NH listeners repeated the sensory dissonance judgment task at lower sensation levels for both sets of PT dyads and the set of HC dyads. All dyads were presented at 53 dB SPL, 30 dB lower than the original presentation level of 83 dB SPL. The levels of the individual components of the dyads were 50 and 42 dB SPL for the PT and HC dyads, respectively. At these SPLs, the sensation levels of the stimuli for these NH listeners were approximately equal to the original sensation levels for the HI listeners, i.e., approximately 20–50 dB SL for the PT dyad components and approximately 10–40 dB SL for the HC

dyad components. The resulting median pure-tone and harmonic complex sensory dissonance curves were very similar in shape and range of scores to those obtained by the NH group at the higher intensity level. This finding suggests that the loss of contrast in the HI group resulted from characteristics of the hearing impairment and not the lower sensation level of the stimuli.

D. Relationship to music perception and training

The data reported here were obtained in a laboratory setting using isolated, artificial stimuli lacking a broader musical context. Under such conditions, judgments of dissonance are made largely on the basis of perceived roughness (Terhardt, 1974). In evaluating the quality of music, however, the unpleasantness due to roughness may be mitigated by other factors. For example, typical musical sounds include an attack portion in which the amplitudes of the partials change rapidly over time. Often, the spectra are highly complex, as when several harmonic complex tones are sounded simultaneously in a chord. Certain pitches may be produced by more than one voice, giving a chorus effect, or sounds may be frequency- or amplitude-modulated, as in vibrato. The effects of these factors on dissonance perception were not assessed in this study. Each, however, would likely reduce the contribution of sensory dissonance to the overall quality of music listening for a NH person (Terhardt, 1978). It is not known how these factors would impact dissonance perception by people with SNHL.

None of the subjects was trained in the sensory dissonance judgment task prior to data collection. The need for practice was judged to be minimal for several reasons. The task was subjective, with no "correct" answer, and each dyad was heard 50 times throughout the experimental sessions. The subjects' judgments were reasonably consistent, producing interpretable patterns of the data. In addition, the two listeners who participated in the lower-level repetition of the dissonance task produced data nearly identical to the first data set.

Trained musicians may base judgments of dissonance primarily on their knowledge of musical intervals, rather than on purely sensory qualities (Plomp and Levelt, 1965). The present study was designed to investigate differences in the perception of sensory dissonance between NH and HI listeners. Musical training may have obscured these differences. Therefore, subjects were excluded from participation if they reported such training. However, all of the subjects had been exposed to Western music over their lifetimes. It is not known whether and how this informal exposure may have influenced their judgments.

With regard to music listening through hearing aids, a future goal of advanced signal-processing algorithms may be the restoration of the normal contrast between consonance and dissonance for HI listeners. Tramo *et al.* (2001) showed that consonant and dissonant intervals produce very distinctive patterns of activity in the auditory nerve. If such patterns of neural activity are dependent upon a normal or near-normal representation of the signal at the level of the cochlea, then the signal must be altered externally to compen-

sate for the effects of the hearing impairment on the internal representation. One possibility may lie in manipulating the phase spectra of musical signals. This approach may be justified in light of evidence that the phase characteristic of the basilar membrane is altered in sensorineural hearing impairment (e.g., Lentz and Leek, 1999; Oxenham and Dau, 2004).

V. CONCLUSIONS

Judgments of the sensory dissonance of PT and HC dyads by the HI listeners were consistent in some respects with those of the NH listeners in this and previous studies. However, several differences were observed. HI listeners did not judge the unison to be as consonant relative to other dyads as the NH listeners did. For the HC dyads, NH listeners judged the musically significant intervals of the unison, octave, fifth, and fourth to be very consonant; HI listeners also judged the unison, octave, and the fifth to be very consonant, but did not clearly judge the fourth to be consonant relative to neighboring intervals. NH listeners showed regions of marked dissonance near the unison, octave, and fifth; HI listeners had a region of marked dissonance near the unison only. These findings suggest that the HI listeners did not distinguish the relative sensory dissonance of intervals as clearly as the NH listeners did. By extension, they may not fully experience the variations in musical tension supplied by dissonant and consonant intervals. The loss of contrast may have resulted from distortions in the representation of pitch in the impaired auditory system (e.g., a reduction in pitch strength), or from more extensive interactions among the components of the dyads. Judgments of sensory dissonance by NH and HI listeners were roughly consistent with a relationship between peripheral frequency selectivity and dissonance perception. A future goal of advanced signal-processing algorithms might be the restoration of the normal contrast between consonant and dissonant intervals for HI listeners, perhaps by altering the phase spectra of musical signals.

ACKNOWLEDGMENTS

This research was supported by a grant from NIH-NIDCD (No. DC 00626). It was approved by the Clinical Investigation Committee and the Human Use Committee, Department of Clinical Investigation, Walter Reed Army Medical Center, under Work Unit No. 03-25012. All subjects participating in this research provided written informed consent prior to beginning the study. The authors would like to thank Robert Lutfi and two anonymous reviewers for their comments on an earlier version of this article, as well as the staff of the Research Section of the Army Audiology and Speech Center at Walter Reed Army Medical Center for their helpful suggestions and discussions. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

¹HC dyads forming a unison or an octave had overlapping components at six and three frequencies, respectively. The two components of the PT dyad forming a unison were of identical frequency. Since the phases of the dyad

components were randomly selected, it is possible that partial or complete cancellation of the overlapping frequency components could have occurred in some cases. The amplitudes of all of the dyads were normalized prior to D/A conversion, however, so that the randomly chosen phases did not result in changes in the levels of the dyads as presented to the subjects.

²In order to facilitate curve fits, the normalized scores were expressed as a function of the frequency separation between dyad components divided by their geometric mean frequency (either 500 or 2000 Hz), or $(f_2 - f_1) / \sqrt{(f_1 * f_2)}$, where f_1 and f_2 were the frequencies of the two pure-tone components, $f_1 < f_2$. Lognormal functions were then fit to the dissonance scores. The abscissa values were converted back to frequency ratios to allow for easier interpretation of the data with regard to musical intervals.

- Agresti, A. (1990). *Categorical Data Analysis* (Wiley, New York).
- Akaike, H. (1974). "A new look at the statistical model identification," *IEEE Trans. Autom. Control* **19**, 716–723.
- American National Standards Institute (ANSI) (1996). "American National Standard: Specifications for audiometers," ANSI S3.6-1996.
- Arehart, K. H., and Burns, E. M. (1999). "A comparison of monotic and dichotic complex-tone pitch perception in listeners with hearing loss," *J. Acoust. Soc. Am.* **106**, 993–997.
- Bacon, S. P., and Gleitman, R. M. (1992). "Modulation detection in subjects with relatively flat hearing losses," *J. Speech Hear. Res.* **35**, 642–653.
- Baker, R. J., and Rosen, S. (2002). "Auditory filter nonlinearity in mild/moderate hearing impairment," *J. Acoust. Soc. Am.* **111**, 1330–1339.
- Bradley, R. A., and Terry, M. E. (1952). "The rank analysis of incomplete block designs. I. The method of paired comparisons," *Biometrika* **39**, 324–345.
- Burns, E. M., and Turner, C. (1986). "Pure-tone pitch anomalies. II. Pitch-intensity effects and diplacusis in impaired ears," *J. Acoust. Soc. Am.* **79**, 1530–1540.
- Chasin, M. (2003). "Music and hearing aids," *Hearing J.* **56**, 36–41.
- David, H. A. (1988). *The Method of Paired Comparisons*, 2nd ed. (Griffin, London).
- deLaat, J. A. P. M., and Plomp, R. (1985). "The effect of competing melodies on melody recognition by hearing-impaired and normal-hearing listeners," *J. Acoust. Soc. Am.* **78**, 1574–1577.
- Florentine, M., Buus, S., and Geng, W. (2000). "Toward a clinical procedure for narrowband gap detection. I. A psychophysical procedure," *Audiology* **39**, 161–167.
- Gfeller, K., Christ, A., Knutson, J. F., Witt, S., Murray, K. T., and Tyler, R. (2000). "Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients," *J. Am. Acad. Audiol.* **11**, 390–406.
- Glasberg, B. R., and Moore, B. C. J. (1986). "Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments," *J. Acoust. Soc. Am.* **79**, 1020–1033.
- Glasberg, B. R., and Moore, B. C. J. (1990). "Derivation of auditory filter shapes from notched-noise data," *Hear. Res.* **47**, 103–138.
- Green, D. M. (1993). "A maximum-likelihood method for estimating thresholds in a yes-no task," *J. Acoust. Soc. Am.* **93**, 2096–2105.
- Greenwood, D. D. (1991). "Critical bandwidth and consonance in relation to cochlear frequency-position coordinates," *Hear. Res.* **54**, 164–208.
- Gu, X., and Green, D. M. (1994). "Further studies of a maximum-likelihood yes-no procedure," *J. Acoust. Soc. Am.* **96**, 93–101.
- He, N., Dubno, J. R., and Mills, J. H. (1998). "Frequency and intensity discrimination measured in a maximum-likelihood procedure from young and aged normal-hearing subjects," *J. Acoust. Soc. Am.* **103**, 553–565.
- Huron, D. (2001). "Tone and voice: A derivation of the rules of voice-leading from perceptual principles," *Music Percept.* **19**, 1–64.
- Hutchinson, W., and Knopoff, L. (1978). "The acoustic component of Western consonance," *Interface (USA)* **7**, 1–29.
- Kameoka, A., and Kuriyagawa, M. (1969a). "Consonance theory. I. Consonance of dyads," *J. Acoust. Soc. Am.* **45**, 1451–1459.
- Kameoka, A., and Kuriyagawa, M. (1969b). "Consonance theory. II. Consonance of complex tones and its calculation method," *J. Acoust. Soc. Am.* **45**, 1460–1469.
- Larkin, W. D. (1983). "Pitch vulnerability in sensorineural hearing impairment," *Audiology* **22**, 480–493.
- Leek, M. R., Dubno, J. R., He, N., and Ahlstrom, J. B. (2000). "Experience with a yes-no single-interval maximum-likelihood procedure," *J. Acoust. Soc. Am.* **107**, 2674–2684.
- Leek, M. R., and Summers, V. (1993). "Auditory filter shapes of normal-hearing and hearing-impaired listeners in continuous broadband noise," *J. Acoust. Soc. Am.* **94**, 3127–3137.

- Leek, M. R., and Summers, V. (2001). "Pitch strength and pitch dominance of iterated rippled noise in hearing-impaired listeners," *J. Acoust. Soc. Am.* **109**, 2944–2954.
- Lentz, J. J., and Leek, M. R. (1999). "Masking by harmonic complexes with different phase spectra in hearing-impaired listeners," *J. Acoust. Soc. Am.* **106**, 2146.
- Moore, B. C. J. (2001). "Dead regions in the cochlea: Diagnosis, perceptual consequences, and implications for the fitting of hearing aids," *Trends in Amplification* **5**, 1–34.
- Oxenham, A. J., and Dau, T. (2004). "Masker phase effects in normal-hearing and hearing-impaired listeners: Evidence for peripheral compression at low signal frequencies," *J. Acoust. Soc. Am.* **116**, 2248–2257.
- Patterson, R. D., and Moore, B. C. J. (1986). "Auditory filters and excitation patterns as representations of frequency resolution," in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London), pp. 123–127.
- Peters, R. W., and Moore, B. C. J. (1992). "Auditory filter shapes at low center frequencies in young and elderly hearing-impaired subjects," *J. Acoust. Soc. Am.* **91**, 256–266.
- Plomp, R., and Levelt, W. J. M. (1965). "Tonal consonance and critical bandwidth," *J. Acoust. Soc. Am.* **38**, 548–560.
- Plomp, R., and Steeneken, H. J. M. (1968). "Interference between two simple tones," *J. Acoust. Soc. Am.* **43**, 883–884.
- Pressnitzer, D., and McAdams, S. (1999). "Two phase effects in roughness perception," *J. Acoust. Soc. Am.* **105**, 2773–2782.
- Rosen, S., and Baker, R. J. (1994). "Characterizing auditory filter nonlinearity," *Hear. Res.* **73**, 231–243.
- Terhardt, E. (1974). "Pitch, consonance, and harmony," *J. Acoust. Soc. Am.* **55**, 1061–1069.
- Terhardt, E. (1978). "Psychoacoustic evaluation of musical sounds," *Percept. Psychophys.* **23**, 483–492.
- Terhardt, E. (1984). "The concept of musical consonance: A link between music and psychoacoustics," *Music Percept.* **1**, 276–295.
- Tramo, M. J., Cariani, P. A., Delgutte, B., and Braida, L. D. (2001). "Neurobiological foundations for the theory of harmony in Western tonal music," *Ann. N.Y. Acad. Sci.* **930**, 92–116.
- Uppenkamp, S., Fobel, S., and Patterson, R. D. (2001). "The effects of temporal asymmetry on the detection and perception of short chirps," *Hear. Res.* **158**, 71–83.
- von Helmholtz, H. (1877/1954). *On the Sensations of Tone as a Physiological Basis for the Theory of Music* (Dover, New York).
- Vos, J. (1988). "Subjective acceptability of various regular twelve-tone tuning systems in two-part musical fragments," *J. Acoust. Soc. Am.* **83**, 2383–2392.
- Zwicker, E., and Fastl, H. (1990). *Psychoacoustics: Facts and Models* (Springer, New York).