Technical University of Denmark

Temporal integration of loudness, loudness discrimination, and the form of the loudness function

Buus, Søren; Florentine, Mary; Poulsen, Torben

Published in:
Acoustical Society of America. Journal

Link to article, DOI:
10.1121/1.417959

Publication date:
1997

Document Version
Publisher's PDF, also known as Version of record

Temporal integration of loudness, loudness discrimination, and the form of the loudness function[a),b)]

Søren Buus
Communication and Digital Signal Processing Center, Department of Electrical and Computer Engineering (409DA), Northeastern University, Boston, Massachusetts 02115

Mary Florentine
Communication Research Laboratory, Department of Speech-Language Pathology and Audiology (133FR), Northeastern University, Boston, Massachusetts 02115

Torben Poulsen
Department of Acoustic Technology (Building 352), Technical University of Denmark, DK-2800 Lyngby, Denmark

(Received 7 May 1996; accepted for publication 25 July 1996)

Temporal integration for loudness of 5-kHz tones was measured as a function of level between 2 and 60 dB SL. Absolute thresholds and levels required to produce equal loudness were measured for 2-, 10-, 50-, and 250-ms tones using adaptive, two-interval, two-alternative forced-choice procedures. The procedure for loudness balances was new and employed ten interleaved tracks to obtain concurrent measurements for ten tone pairs. Each track converged at the level required to make the variable stimulus just louder than the fixed stimulus. Thus, the data yield estimates of the just-noticeable difference (jnd) for loudness level and temporal integration for loudness. Results for four listeners show that the amount of temporal integration, defined as the level difference between equally loud short and long tones, varies markedly with level and is largest at moderate levels. The effect of level increases as the duration of the short stimulus decreases and is largest for comparisons between the 2- and 250-ms tones. The loudness-level jnds are also largest at moderate levels and, contrary to traditional jnds for the level of two equal-duration tones, they do not appear to depend on duration. The latter finding indicates that loudness discrimination between stimuli that differ along multiple dimensions is not the same as level discrimination between stimuli that differ only in level. An equal-loudness-ratio model, which assumes that the ratio of loudnesses for a long and short tone at equal SPL is the same at all SPLs, can explain the level dependence of temporal integration and the loudness jnds. It indicates that the loudness function [log(loudness) versus SPL] is flatter at moderate levels than at low and high levels in agreement with earlier findings for 1-kHz tones [M. Florentine *et al.*, J. Acoust. Soc. Am. **99**, 1633–1644 (1996)]. © *1997 Acoustical Society of America.* [S0001-4966(97)02112-7]

PACS numbers: 43.66.Cb, 43.66.Mk, 43.66.Ba [WJ]

INTRODUCTION

The purpose of this paper is to examine the effect of level on the amount of temporal integration for loudness, defined as the level difference between equally loud short and long sounds. This information is important to understand the perception of natural sounds and may also provide insight into intensity coding in the auditory system.

Temporal integration of loudness for tones has been measured at more than one level in a number of studies (e.g., Munson, 1947; Ekman *et al.*, 1966; Stephens, 1973; Pedersen *et al.*, 1977; Poulsen, 1981; Florentine *et al.*, 1996). However, it has been difficult to ascertain how the amount of temporal integration depends on level, because most studies examined only a limited range of levels and many did not use the very short durations that are likely to yield the largest effect of level. To our knowledge, only two studies measured temporal integration of loudness of tones over wide ranges of levels and durations (Munson, 1947; Florentine *et al.*, 1996). Unfortunately, Munson (1947) used a method in which a test tone (5, 10, 40, 100, or 200 ms) always preceded a 1-s tone and only the 1-s tone was varied in level. Both of these procedural features are known sources of bias. For example, the second of two identical tones tends to sound louder than the first (e.g., Stevens and Davis, 1938) and listeners tend to underestimate the loudness of the variable stimulus at low levels and overestimate it at high levels (e.g., Stevens, 1956; Stevens and Greenbaum, 1966; Florentine *et al.*, 1996). In addition, it was not possible to determine the internal consistency of Munson's loudness matches because no control conditions were included in the measurements. These shortcomings make the data difficult to interpret.

[a)]Parts of this paper were presented at the 11th Meeting of the International Society of Psychophysics, October 1995, Cassis, France [S. Buus *et al.*, in *Fechner Day '95*, edited by C.-A. Possamaï (Int. Soc. Psychophysics, Cassis, France), pp. 55–60 (1995)] and at the 131st meeting of the Acoustical Society of America, May 1996, Indianapolis, Indiana [S. Buus *et al.*, J. Acoust. Soc. Am. **99**, 2490 (A) (1996)].

[b)]''Selected research articles'' are ones chosen occasionally by the Editor-in-Chief, that are judged (a) to have a subject of wide acoustical interest, and (b) to be written for understanding by broad acoustical readership.

In contrast, Florentine *et al.* (1996) used a modern psychophysical procedure and included control conditions. These data indicated that the amount of temporal integration for loudness of 1-kHz tones varied nonmonotonically with level. The effect of level was greatest at short durations. Therefore, one purpose of the present study was to examine the effect of level on temporal integration of loudness with durations even shorter than those used in our previous study. To minimize confounding effects of spectral splatter that arise when the stimuli are brief, the present study used tones at 5 kHz where the critical band is wide enough to accommodate very brief tone bursts.

Another purpose of the present study was to investigate a new loudness-balance procedure that was designed to minimize biases and increase consistency. Although Florentine *et al.*'s (1996) data showed a high degree of internal consistency, they indicated a small bias that caused the variable stimulus to migrate towards a comfortable loudness. In addition, the judgments were not completely transitive, because the level difference between equally loud 5- and 200-ms stimuli tended to be larger than the sum of level differences between 5- and 30-ms stimuli and between 30- and 200-ms stimuli. Thus, our new procedure attempted to minimize biases by interleaving adaptive tracks for several conditions such that listeners would not know which stimulus was the variable in a particular trial. Control conditions were included to allow an evaluation of transitivity in the data.

Because the new procedure was designed to estimate the level at which the variable-level tone was just louder than the fixed-level tone, it also provided measurements of difference limens for loudness level as a function of level and duration of the variable tone (cf. Schlauch and Wier, 1987; Zeng, 1994). Thus, a third purpose of the present study was to compare difference limens for loudness level to DLs obtained in traditional level-discrimination experiments in which the listeners compare two tones that have the same duration and differ only in level (e.g., Florentine, 1986; Buus and Florentine, 1992). As discussed later, the comparison of difference limens obtained for tones that differ both in duration and level with those for equal-duration tones differing only in level may reveal important differences between loudness discrimination and level discrimination.[1]

I. METHOD

A. Stimuli

The stimuli were 5-kHz tone bursts with nominal durations of 2, 10, 50, and 250 ms. These durations were measured between the half-amplitude points of the 2-ms linear rise and fall. The equivalent rectangular durations are 0.67 ms shorter than the nominal durations. The 2-ms stimulus consisted of only the rise and fall. The longer stimuli contained steady-state segments with a duration 2 ms less than the nominal duration. These envelope shapes ensured that almost all the energy of the tone bursts was contained within the 900-Hz-wide critical band centered at 5 kHz (cf. Zwicker, 1961; Scharf, 1970). Even for the 2-ms tone burst, the energy within the critical band was only 0.02 dB less than the overall energy. Nine test levels were used for each duration of the fixed-level stimulus. They were 2, 5, 10, 15, 20, 30, 40, 50, and 60 dB SL. All SPLs reported are equivalent free-field SPLs as determined by threshold measurements in free field and with the Sony MDR-V6 earphone in a group of ten normal listeners (cf. Villchur, 1969).

B. Procedure

1. Absolute thresholds

In the first part of the experiment, absolute thresholds were measured for each of the four test durations using an adaptive procedure in a two-interval, two-alternative forced-choice (2I, 2AFC) paradigm. Each trial contained two observation intervals, marked by lights and separated by 500 ms. The signal was presented in either the first or the second observation interval with equal *a priori* probability. The listener's task was to indicate which interval contained the signal by pressing a key. One hundred milliseconds after the listener responded, the correct answer was indicated by a 200-ms light. Following the feedback, the next trial began after a 200-ms delay.

A single threshold measurement was based on three interleaved adaptive tracks, each of which ended after five reversals. Reversals occurred when the signal level changed from increasing to decreasing or *vice versa*. On each trial, the track was selected at random among the tracks that had not yet ended. For each track, the level of the signal initially was set approximately 15 dB above the listener's threshold. It decreased following three correct responses and increased following one incorrect response, such that the signal converged on the level yielding 79.4% correct responses (Levitt, 1971). The step size was 5 dB until the second reversal, after which it was 2 dB.

The threshold for one track was calculated as the average signal level at the fourth and fifth reversals, and one threshold measurement was taken as the average threshold across the three tracks. Three such threshold measurements (for a total of nine tracks) were obtained for each listener and condition. The average across all measurements was used as the reference to set the sensation level for each listener and condition in the second part of the experiment.

2. Loudness matches

In the second part of the experiment, stimuli of different durations were matched in loudness to one another using a new procedure, which combines features of Fletcher and Munson's (1933) forced-choice procedure and Jesteadt's (1980) adaptive procedure for loudness matching. The new procedure was intended to eliminate some biases that may have affected previous loudness-matching data obtained with adaptive procedures. For example, if all trials in a measurement use the same pair of sounds (except for small changes in level of the variable sound), listeners may start to ignore the fixed sound, which is always the same, and base their judgments on comparisons with previous trials. If the variable sound in a trial sounds louder than it did in the previous trial, the listener is likely to respond ''variable louder.'' In addition, judgments that ignore the fixed sound are likely to be affected by a comparison to comfortable loudness such that the variable is more likely to be judged ''too soft'' when the sounds are soft and ''too loud'' when they are loud (Florentine *et al.*, 1996).

To reduce such biases, the present experiment employed a procedure in which the loudnesses of both the long and the short tones vary from trial to trial. This is similar to the random selection of stimulus pairs for each trial used by Fletcher and Munson (1933) for loudness matching with the method of constant stimuli. The random variation of overall loudness level was intended to force the listeners to base their responses on loudness judgments of the two tones presented in the trial. Roving loudness and mixing trials with both the long and the short tone varied also help listeners focus on loudness and ignore other differences, which remain constant throughout a measurement.

Loudness matches were made at nine levels for each of five duration pairs (i.e., 2 vs 10 ms, 2 vs 250 ms, 10 vs 50 ms, 10 vs 250 ms, and 50 vs 250 ms). To keep the number of trials within a block of trials reasonable, the nine levels were divided into a low-SL set (i.e., 2, 5, 10, 15, and 20 dB SL) and a high-SL set (i.e., 20, 30, 40, 50, and 60 dB SL). Both sets included 20 dB SL, which allowed the possible effect of loudness range on listeners' judgments to be examined. For each listener and duration pair, three low-SL and three high-SL sets were obtained in mixed order. Each set consisted of ten interleaved tracks for one duration pair: matches were made at five different levels with both the long and the short tones varied. A set was completed in about 250 trials.

The loudness matches were obtained in a 2I, 2AFC paradigm. On each trial, the listener heard two tones separated by 600 ms. The fixed-level tone followed the variable tone or the reverse with equal *a priori* probability. The listener's task was to indicate which sound was louder by pressing a key. The response initiated the next trial after a 1-s delay. The level of the variable tone for each track was adjusted according to a one-up, two-down procedure. If the listener indicated that the variable tone was louder than the fixed tone on two successive trials for a particular track, its level was reduced, otherwise it was increased. The step size was 5 dB until the second reversal, after which it was 2 dB.

For each track, the variable stimulus was initially set approximately 15 dB below the expected equal-loudness level. (If that level was below threshold, the variable stimulus was set to its threshold.) The choice of starting levels ensured that the listener would hear trials in which both the short and the long tone was clearly the louder one, because both the long and short tones were varied within a block of trials. On each trial, the track was chosen at random among those that had not yet ended, which they did after nine reversals. The average level of the last four reversals of each track was used as an estimate of the level at which the variable tone was just louder than the fixed-level tone. This procedure made the variable tone converge towards the level at which it was judged louder than the fixed tone in 71% of the trials (Levitt, 1971). Thus, the procedure provides an estimate of the level at which the variable is just noticeably louder than the fixed-level tone (Schlauch and Wier, 1987). These ''just-louder'' levels are used in the subsequent data analysis.

C. Apparatus

An APR 486/33 PC-compatible computer controlled the stimulus generation, sampled the listeners' responses, and executed the adaptive procedure. The tone bursts were produced by a programmable waveform generator (TDT WG1) whose output was attenuated (two TDT PA4s in series), and led to a headphone amplifier (TDT HB5), which fed one earpiece of a Sony MDR-V6 headset.

The earphone was calibrated daily on an acoustic test fixture (ISO, TR 4869-3, 1989). The microphone (Brüel & Kjær 4144) received its DC bias from a measuring amplifier (Brüel & Kjær 2607), which also measured the SPL produced by the earphone.

D. Listeners

Four listeners, three males and one female, were tested on all conditions. One relatively inexperienced listener, SP, was 25 years old and the other listeners, who are the authors, were between 41 and 49. The authors had extensive experience making loudness judgments; SP practiced loudness judgments for one 2-h session before data collection. The listeners had normal audiometric thresholds (ISO 389, 1991) at standard octave frequencies from 0.125–8 kHz, except for one author who had a mild cochlear hearing loss below 2 kHz. Because that listener's data were similar to those of the others, data from all listeners were averaged.

E. Data analysis

For each listener and condition, the three ''just-louder'' levels were averaged. These averages were then used to calculate group averages and standard deviations across listeners. Because our primary interest was to investigate the effect of level on the amount of temporal integration, statistical analyses were performed on the level differences between the short and the long tone when the variable was judged ''just louder.'' To examine the statistical significance of the effects of stimulus variables and differences among listeners, a four-way ANOVA (stimulus level×duration pair×long or short variable×listener) for repeated measures was performed (Data Desk 4.2, Data Description, Inc., Ithaca, NY, 1994). Scheffé *post hoc* tests for contrast (Data Desk 4.2, 1994) were performed as appropriate to explore the sources of significant effects and interactions. Unless otherwise stated, differences were considered significant when these tests indicated a probability less than 0.05.

II. RESULTS

A. Individual listeners

Figure 1 shows the levels at which each listener judged the variable to be just louder than the fixed stimulus. The unfilled symbols show the ''just-louder'' levels for the short tones; they show the level at which the short tone is judged louder in 71% of the trials, which corresponds to one just-noticeable difference above the equal-loudness level. The filled symbols show the ''just-louder'' levels for the long tones; they show the level at which the long tone is judged louder in 71% of the trials, which corresponds to one just-

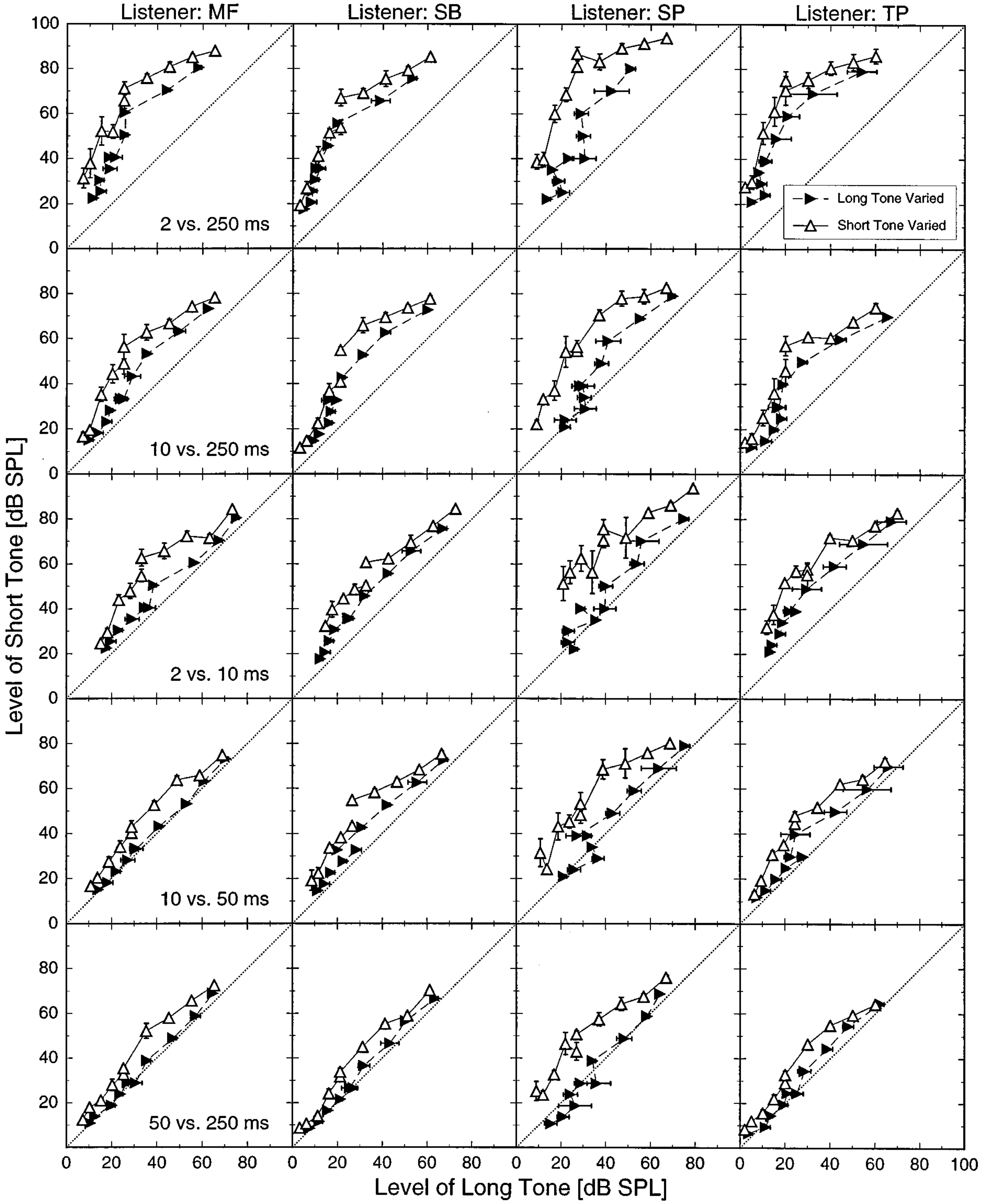

FIG. 1. Individual loudness judgments for four listeners. The level of the short tone is plotted as a function of the level of the long tone. The symbols show the levels of the variable stimuli at which they were judged ''just louder'' than the fixed stimuli. The unfilled symbols show data obtained when the short tones were varied; the filled symbols show data obtained when the long tones were varied. Each row of panels shows the data for a particular duration pair. Each column shows the data for an individual listener. The error bars show plus and minus one standard error calculated across the three repetitions for each listener and condition.

noticeable difference to the right of the equal-loudness level. As indicated by the error bars, the listeners' judgments are quite consistent. The average standard error is 2.3 dB. However, the data for the less experienced listener, SP, tend to be more irregular and have larger standard errors than those for the experienced listeners. Although the ANOVA, which is discussed later, showed some highly significant interlistener differences, the data are reasonably consistent across listeners. The average across-listener standard deviations of the level differences required for the variable stimuli to be just louder is 4.1 dB.

The amount of temporal integration for the data in Fig. 1

is indicated by the distance from the centroid of the two functions to the dotted diagonal, which shows equal levels of the short and long tones. For all four listeners, the distance is larger at moderate levels than at low and high levels, which indicates that the amount of temporal integration is a nonmonotonic function of level in agreement with the findings for 1-kHz tones and noises in our previous study (Florentine *et al.*, 1996). For the duration pairs differing by a factor of 5, the effect of level is relatively modest, but it clearly increases as the tones become briefer: the data for the 2- and 10-ms tones are considerably farther from the diagonal and show a larger effect of level than those for the 50- and 250-ms tones. As the duration ratio increases, both the amount of temporal integration and the effect of level increase. For the 2- and 250-ms tones, the amount of temporal integration is 15–20 dB at low levels, reaches a peak of about 40 dB when the 250-ms tone is around 20 dB SL, and decreases to about 20–25 dB at the highest levels. (Thresholds for the 250-ms tone were 5.3 dB SPL for MF, 1.2 dB SPL for SB, 7.0 dB SPL for SP, and 0.1 dB SPL for TP.)

The just-noticeable differences in loudness level are indicated by the separation between the data for the short tone varied and the long tone varied. This distance tends to be larger for SP than for the other listeners, indicating that SP's just-noticeable differences for loudness level are larger than those for the experienced listeners. For all four listeners, the jnd's appear to be larger at moderate levels than at levels below 25 dB SPL and above 65 dB SPL. This mid-level ''hump'' of the jnds for loudness level is qualitatively similar to the jnds obtained by Zeng (1994), when listeners compared the loudness of a forward-masked tone to that of a tone presented in isolation. Somewhat surprisingly, no consistent difference is apparent among the five duration pairs. Certainly, the filled and unfilled functions for the 50- and 250-ms tone pair are not markedly closer together than those for the 2- and 10-ms or even 2- and 250-ms tone pairs.

Each function contains two data points with the fixed stimulus set to 20 dB SL, because this level was included in both the low- and the high-SL set. The difference between the two points indicates the extent to which the listeners' judgments depended on the range of levels presented in a block of trials. Generally, the level at which the variable is just louder than the 20-dB SL fixed stimulus is higher for the high- than for the low-SL set. The average difference is 2.8 dB. Thus, the data show a context effect, which may be related to the range of loudnesses encompassed by the stimuli tested within the block. This effect is reminiscent of the context effects in Durlach and Braida's (1969) memory model for intensity perception. It should be noted that this context effect generally is much smaller than the amount of temporal integration.

B. Group data

The large symbols in Fig. 2 show the average data for the four listeners plotted in the same manner as Fig. 1. The small symbols and the solid lines show predictions by a model, which is discussed later. The average data reflect the trends in the individual data. The amount of temporal integration varies markedly with level and is largest at moderate levels, especially for the short durations and large duration ratios. The just-noticeable difference in loudness level, indicated by the vertical distance to the solid line for the unfilled symbols and the horizontal distance for the filled symbols, also is largest at moderate levels. A comparison across panels shows no systematic effect of duration on the jnd. Finally, the data at 20 dB SL indicate that the jnd generally is larger for the high-SL set than for the low-SL set. This context effect varies somewhat across durations and tends to be largest when the short tone is varied.

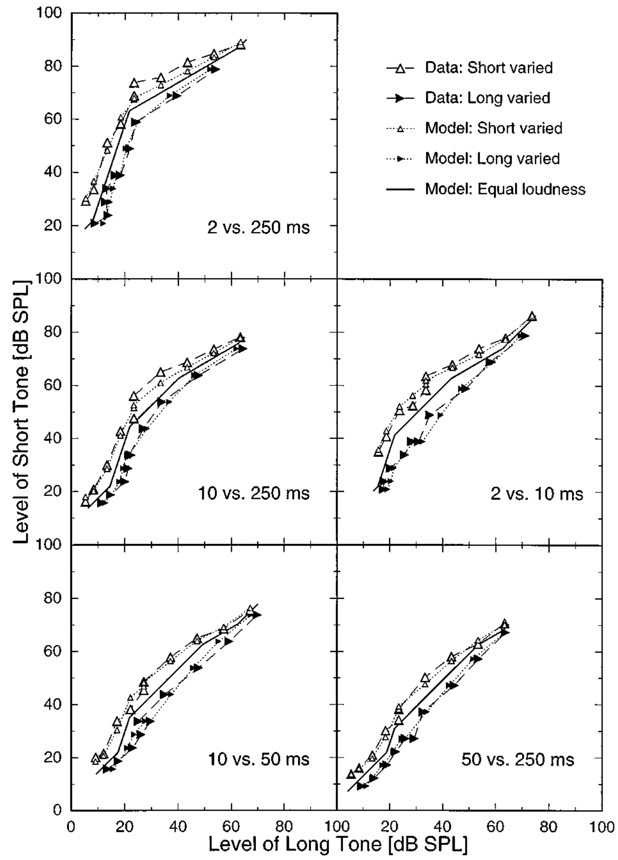

FIG. 2. Average of loudness judgments across four listeners. The level of the short tone is plotted as a function of the level of the long tone. As in Fig. 1, unfilled symbols show the ''just louder'' levels for the short tones and filled symbols show the ''just-louder'' levels for long tones. The large symbols show the data and the small symbols show predictions by a model (see text). The solid lines show equal-loudness functions predicted by the model. Each panel shows data for a different duration pair.

These effects are supported by the ANOVA shown in Table I. Effects of stimulus variables that are consistent across listeners are shown in the top half of the table and individual differences (i.e., effect of and interactions with listener) in the bottom half. The top half of the table shows that all the experimental variables had highly significant effects on the level difference between the short and the long tone. The effect of SL shows that the amount of temporal integration varies with level. The Scheffé tests for contrast showed that, overall, the amount of temporal integration is larger at 20, 30, and 40 dB SL than at lower and higher levels. However, this effect depends both on the duration pair and whether the short or the long tone is varied as indicated by the interactions between SL and these other vari-

TABLE I. Four-way analysis of variance for repeated measures of loudness matching. The dependent variable is the level difference between the variable and the fixed stimulus, when the variable was judged ''just louder'' than the fixed stimulus. The stimulus variables Sensation Level (SL; 9 levels: 2,5,10,15,20,30,...,60 dB SL), Pair (Pr; 5 levels: 2 vs 250 ms, 10 vs 250 ms, 2 vs 10 ms, 10 vs 50 ms, and 50 vs 250 ms), and Variable (Vrb; 2 levels: Short or Long) are fixed factors. Listener (Lsr; 4 levels: MF, SB, SP, and TP) is a random factor.

Source	df	Error df	Sums of squares	Mean square	F ratio	Probability
Const	1	3	273 204	273 204	357.5	0.0003
SL	8	24	17 399	2175	88.70	⩽0.0001
Pr	4	12	66 011	16 503	94.40	⩽0.0001
SL*Pr	32	96	6649	207.8	8.948	⩽0.0001
Vrb	1	3	35 688	35 688	19.01	0.0223
SL*Vrb	8	24	6867	858.4	16.70	⩽0.0001
Pr*Vrb	4	12	473.7	118.4	7.958	0.0023
SL*Pr*Vrb	32	96	2750	85.95	3.321	⩽0.0001
Lsr	3	840	2293	764.2	28.30	⩽0.0001
SL*Lsr	24	840	588.5	24.52	0.908	0.5915
Pr*Lsr	12	840	2098	174.8	6.474	⩽0.0001
SL*Pr*Lsr	96	840	2229	23.22	0.860	0.8239
Vrb*Lsr	3	840	5633	1878	69.53	⩽0.0001
SL*Vrb*Lsr	24	840	1234	51.41	1.904	0.0057
Pr*Vrb*Lsr	12	840	178.6	14.88	0.551	0.8813
SL*Pr*Vrb*Lsr	96	840	2485	25.88	0.958	0.5935
Error	840		22 683.7	27.00		
Total	1199		189 234			

ables. The Scheffé tests showed no significant differences among the various levels for duration pairs that differed only by a factor of 5, but for the 10- and 250-ms and the 2- and 250-ms pairs, the amount of temporal integration was significantly larger between 20 and 50 dB SL than at lower or higher levels.

As expected, the effect of duration pair is highly significant. Although the magnitude of this effect depends on whether the long or the short tone is varied and on the SL, the significant differences were the same whether the long or the short tone was varied. When the tone durations differed by a factor of 5, the amount of temporal integration increased significantly as duration decreased. The Scheffé tests showed that, averaged across SLs, the amount of temporal integration is significantly larger for the 2- and 10-ms pair than for the 10- and 50-ms tones, which yielded more temporal integration than the 50- and 250-ms tones. As expected, the amount of temporal integration increased as the duration ratio increased. The 2- and 250-ms pair yielded more temporal integration than the 10 and 205 ms pair, which yielded more temporal integration than the 50- and 250-ms pair.

Although the effects in the average data generally are apparent in the individual data, the analysis also indicated some highly significant differences among listeners as evident in the lower half of Table I. The amount of temporal integration differs somewhat among listeners. The Scheffé tests show that, overall, the amount of temporal integration is larger for TP than for SB, and smaller for MF than for the other three listeners. The differences are small, however, about 1.6 dB between TP and SB and an average of 3 dB between MF and the other three listeners. These differences depend somewhat on condition as indicated by the significant interactions between listener and duration pair, between listener and variable (long or short), and among listener, stimulus level, and variable.

III. DISCUSSION

A. Model of temporal integration and discrimination of loudness

The results indicate that the just-noticeable loudness-level difference and the amount of temporal integration show similar dependencies on level: both are larger at moderate levels than at low and high levels. This similarity may result from a common underlying factor. Florentine *et al.* (1996) argued that the ratio between the loudness of a long and a short tone at equal SPL was likely to be independent of the SPL. In fact, Stevens and Hall's (1996) data for burst of white noise agree with this contention, which was also a fundamental property of Zwislocki's (1969) analysis of temporal integration of loudness. Accordingly, the amount of temporal integration for a given duration pair should be inversely proportional to the slope of the function relating the logarithm of loudness to SPL. To account for the finding of a maximal amount of temporal integration at moderate levels, Florentine *et al.* (1996) assumed that the loudness function was shallower at moderate than at high and low levels. This idea appears applicable to the present data.

To account for the present measurements of ''just louder'' levels, the just-noticeable loudness-level difference also must be considered. For a given just-noticeable difference in loudness, the just-noticeable loudness-level difference is inversely proportional to the slope of the log(loudness)-vs-SPL function. Thus, a loudness function that accounts for the effect of level on the amount of temporal integration of loudness also is likely to account for the present finding of larger just-noticeable loudness-level differences at moderate than at low and high levels. Indeed, the data in Figs. 1 and 2 indicate rather directly that the logarithm of loudness grows considerably more slowly at moderate levels than at low and high levels. Consider, for example, the comparison of 2- and 250-ms tones (top panel in Fig. 2). As the 250-ms tone increases from 5 to 23 dB SPL, the 2-ms tone must increase from about 19 to about 64 dB SPL to maintain equal loudness. An 18-dB change in level at low SPLs increases the loudness as much as a 45-dB change in level at moderate SPLs. On the other hand, as the 250-ms tone increases another 30 dB—from about 25 to 55 dB SPL—the equally loud 2-ms tone increases only 18 dB—from about 65 to about 83 dB SPL. The data for the other duration pairs show similar trends and all indicate that loudness grows more slowly at moderate than at low and high levels.

To investigate if a loudness function of the form described above can account for the present data, these ideas were incorporated into a model. For simplicity, the loudness, N, of a tone was modeled as three power-function segments. The form of such functions resembles the input–output functions of the basilar membrane described by Yates (1990). As in our previous paper (Florentine *et al.*, 1996), the ratio between the loudnesses of short and long tones of equal SPL was assumed to be independent of the SPL. Thus, the loudness of a tone was calculated as

$$N(I,\mathrm{Dur})=k(\mathrm{Dur})*\begin{cases}\left(\frac{I}{I_0}\right)^a, & I\leqslant I_1,\\ \left(\frac{I_1}{I_0}\right)^{(a-b)}\left(\frac{I}{I_0}\right)^b, & I_1<I\leqslant I_2,\\ \left(\frac{I_1}{I_0}\right)^{(a-b)}\left(\frac{I_2}{I_0}\right)^{(b-c)}\left(\frac{I}{I_0}\right)^c, & I>I_2,\end{cases} \tag{1}$$

where I is the intensity of the tone, Dur is its duration, and I_0 is the reference intensity ($=10^{-12}$ W/m^2). The exponents a, b, and c, the intensity limits (or break points) I_1 and I_2, and k(Dur) are free parameters, which are used to fit the data and to make the fitted loudness function for the 250-ms tones similar to a ''standard'' 5-kHz loudness function. The standard loudness function was obtained by using Zwislocki's (1965) modified power function to calculate the loudness in sones from the loudness level in phons. The loudness levels of the 5-kHz tones were determined from the ISO 226 (1987) equal-loudness-level contours.

The just-noticeable difference in loudness was assumed to be proportional to the square root of the loudness as suggested by Hellman and Hellman (1990).[2] To account for the effect of context seen in the data at 20 dB SL, discrimination was assumed to be corrupted by an additional internal noise, which was proportional to the range of loudnesses encompassed by the stimuli in the set of ten adaptive tracks. Thus, the just-noticeable difference in loudness, ΔN_{DL}, was modeled as

$$\Delta N_{\mathrm{DL}}=\sqrt{\partial^2 N_{\mathrm{fix}}+r^2(N_{\max}-N_{\min})^2}, \tag{2}$$

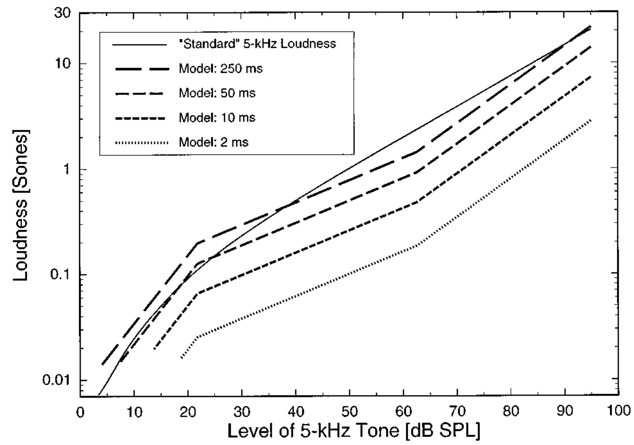

FIG. 3. The thick lines show loudness functions obtained by fitting a simple model to the present data for temporal integration of loudness and loudness-level jnd's. The model loudness function for 250-ms tones is compared to a ''standard'' loudness function for 5-kHz tones (see text for details).

where N_{fix} is the loudness of the fixed stimulus of the pair, $N_{\max}$ and $N_{\min}$ are the maximum and minimum loudnesses of the stimuli in the set of tracks, ∂ is a free parameter that determines the just-noticeable difference for loudness in the absence of roving loudness, and r is a free parameter that determines the magnitude of the context effect.[3]

Accordingly, the intensity of the variable stimulus for a given comparison could be predicted as the intensity, I_{var}, for which

$$N(I_{\mathrm{var}},\mathrm{Dur}_{\mathrm{var}})=N(I_{\mathrm{fix}},\mathrm{Dur}_{\mathrm{fix}})+\Delta N_{\mathrm{DL}}, \tag{3}$$

where I_{fix} is the intensity of the fixed stimulus, $\mathrm{Dur}_{\mathrm{fix}}$ is its duration, and $\mathrm{Dur}_{\mathrm{var}}$ is the duration of the variable stimulus. Optimizing the parameters to yield a loudness function that is similar to the ''standard'' loudness function for 5-kHz tones for the 250-ms duration, while providing a small rms error of the predictions, yielded $a=0.64$, $b=0.21$, $c=0.36$, $k(250\ \mathrm{ms})=0.0077$, $k(50\ \mathrm{ms})=0.0049$, $k(10\ \mathrm{ms})=0.0026$, and $k(2\ \mathrm{ms})=0.0010$. The break points, I_1 and I_2, corresponded to 22 and 63 dB SPL, ∂^2 was 0.021 sones, and r was 0.029.

As shown in Fig. 2, the model predictions (small symbols) obtained with these parameters agree very well with the average data and the predicted equal-loudness functions describe the central tendencies of the data. In addition, no systematic errors are apparent, except that the context effect at 20 dB SPL generally is larger than predicted. Despite the underprediction of the context effect, the rms error is only 2.1 dB. This error is slightly smaller than the average standard error of the individual listeners' data. The model also is able to predict the individual listeners' data with rms errors that approach the rms values of the standard errors for the means across repetitions. The rms error of the model predictions is about 50% larger than the rms value of the standard error for MF and SB, whereas they are about equal for TP and SP. Thus, this simple model provides a satisfactory account of the data, as is also evident in Fig. 2.

The loudness functions obtained from the model are shown by the thick lines in Fig. 3. The model loudness func-

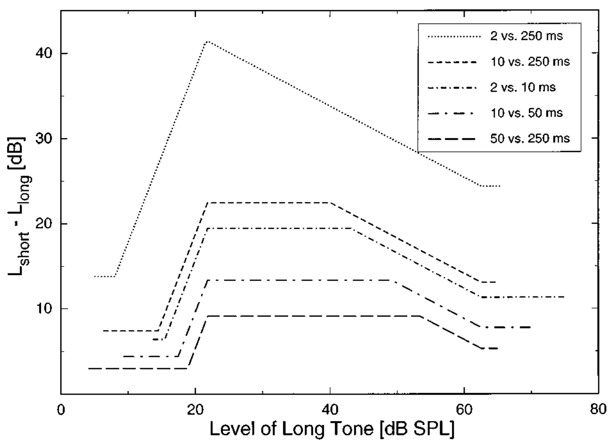

FIG. 4. The amount of temporal integration, defined as the level difference between equally loud short and long tones, is plotted as a function of level of the longer tone. The different lines show results for different duration pairs as indicated by the legend.

tions for the shorter tones are parallel to that for the 250-ms tone in accordance with Florentine *et al.*'s (1996) assumption that the ratio between the loudnesses of a long and a short tone at equal SPL is independent of the SPL. The loudness function for the 250-ms tones is compared to the loudness function obtained by combining Zwislocki's (1965) modified power function and the ISO 226 (1987) equal-loudness contours. The differences between the modified power function and the model loudness function are relatively minor, which indicates that the model loudness function for the 250-ms tones may be considered consistent with existing data. Although the ISO 226 (1987) equal-loudness contours indicate that the loudness of high-level tones grows faster at 5 than at 1 kHz, it should be noted that few data exist to support the relatively steep slope of 0.36 obtained at high levels.

The equal-loudness functions shown in Fig. 2 are derived from the loudness functions in Fig. 3. Because the loudness function was fixed for each stimulus duration, the predicted equal-loudness functions are perfectly transitive. The good agreement between the data and the model predictions indicate that the listeners' judgments also show nearly perfect transitivity. Thus, it appears that our new method for loudness-balance measurements was successful in reducing biases that interfered with transitivity in previous loudness-matching data obtained with adaptive procedures.

B. Effect of level on the amount of temporal integration

The amount of temporal integration clearly varies with level and is largest at moderate levels, as shown in Fig. 4. These functions are obtained from the equal-loudness functions predicted by the model. At low levels, the amount of temporal integration is close to the threshold difference between the stimuli in a pair. (Average thresholds were 3.4, 7.2, 13.7, and 18.8 dB SPL for durations of 250, 50, 10, and 2 ms, respectively.) At moderate levels, it reaches a plateau of three times that obtained at low levels, except for the 2- vs 250-ms comparison. For this pair, the amount of temporal integration reaches a peak around 24 dB SPL; this maximum is also about three times that obtained at low levels. At levels above the plateau or peak, the amount of temporal integration decreases again to reach a lower plateau of about 1.8 times that obtained near threshold.

The amounts of temporal integration obtained in the present study are larger and show more variation with level than those obtained in our previous study for 1-kHz tones and broadband noises (Florentine *et al.*, 1996). This is particularly true for the largest duration ratio used. The amount of temporal integration between 2- and 250-ms tones varies from about 14 dB near threshold to over 40 dB around 20 dB SL; above this level it decreases to reach about 25 dB around 60 dB SL. Thus, the effect of level on temporal integration exceeds 25 dB when the brief tone lasts only 2 ms. This large effect of level clearly exceeds the 8-to-9-dB level effect obtained for 5- and 200-ms stimuli in our previous study. As expected, decreasing the duration of the briefest stimuli increased the effect of level as well as the amount of temporal integration.

The differences between the two studies may not be attributable entirely to the stimulus durations. Even when the duration ratios are roughly comparable, the amounts of temporal integration obtained in the present study exceed those obtained in our previous study. For example, Florentine *et al.* (1996) reported a maximal level difference of about 7 dB between equally loud 30- and 200-ms tones (ratio: 6.7), whereas the present data show a maximum of about 9 dB difference between 50- and 250-ms tones (ratio: 5). For 5- and 200-ms tones (ratio: 40), our previous study obtained a maximal amount of temporal integration of about 19 dB, whereas the present data show about 22 dB between 10- and 250-ms tones (ratio: 25). Although the differences are not large, they are surprising because their direction is opposite what might be expected. The amount of temporal integration increases as the duration ratio increases and the duration decreases. Because, in these comparisons, the duration ratios are smaller and the tones longer in the present study than in the previous study, one would expect the amounts of temporal integration to be smaller in the present study than in the previous study, not larger.

The discrepancy between the present study and our previous study may result from the different frequencies used. The larger amounts of temporal integration obtained in the present study may reflect that the auditory system shows more mid-level compression at high than at low frequencies. Cooper and Yates (1994) showed that input–output functions derived from comparisons between rate-intensity functions for below-CF tones and tones at CF are much more compressive for CFs above 4 kHz than below this frequency. Indeed, the flattest parts the loudness functions derived from our temporal-integration data have exponents of about 0.25 at 1 kHz and 0.21 at 5 kHz. If the loudness ratio between a short and a long tone of equal level is independent of level, a flattening of the loudness function increases the amount of temporal integration. Thus, part of the difference between our previous data for 1-kHz tones and the present data for 5-kHz tones may reflect that the loudness function is more compressive at 5 than at 1 kHz. It should be noted, however,

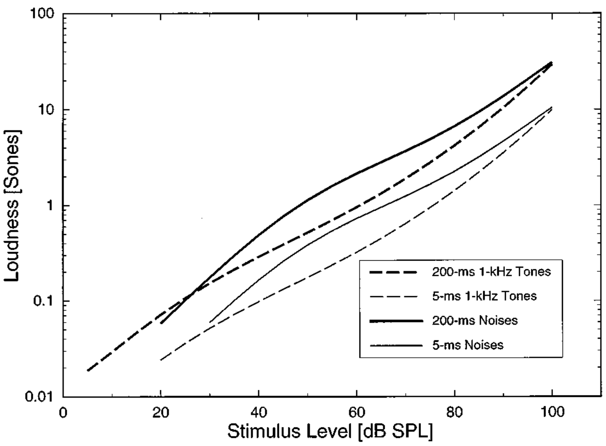

FIG. 5. Loudness functions for 1-kHz tones and white noises derived from Florentine *et al.*'s (1996) data for temporal integration. The dashed lines show loudness functions for 1-kHz tones and the solid lines show loudness functions for white noises. The thick lines are for 200-ms stimuli and the thin lines for 5-ms stimuli.

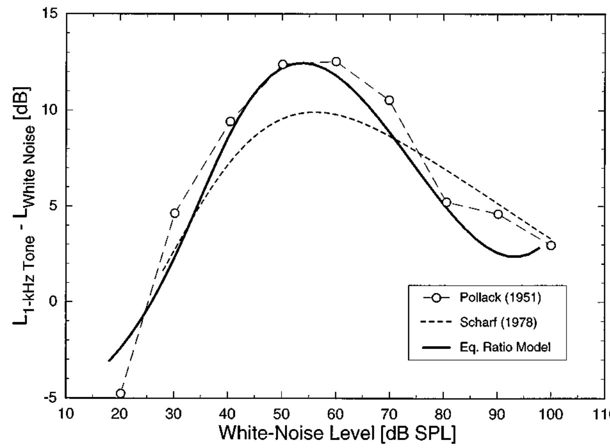

FIG. 6. Loudness summation as a function of level. The level difference between equally loud tones and white noises is plotted as a function of the level of the white noise. The circles show data from Pollack (1951), the dashed line shows data derived from Scharf's (1978) loudness functions for 1-kHz tones and white noises, and the thick solid line shows predictions from the loudness functions in Fig. 5.

that the difference in the compressiveness of the loudness functions at 1 and 5 kHz is much smaller than the difference between the input–output functions reported by Cooper and Yates (1994).

C. Further support for the equal-loudness-ratio model

The hypothesis that the loudness ratio between a short and a long tone of equal level is independent of level may be tested further by using the loudness functions it produces to predict loudness data other than those used to derive them. For example, one may assume that the relative loudnesses of the brief noises in Florentine *et al.*'s (1996) experiment should be the same as those they obtained for the 1-kHz tone. This assumption permits deriving loudness functions for broadband noises based only on the amounts of temporal integration measured for noise bursts, except for a single multiplicative factor. The resulting loudness functions for 5- and 200-ms noise bursts are shown in Fig. 5 together with the loudness functions Florentine *et al.* (1996) derived from their data for 1-kHz tones.

These loudness functions may be used to predict loudness-balance functions between a 1-kHz tone and a wideband noise as shown in Fig. 6. The multiplicative factor allows the loudness functions for the noises in Fig. 5 to be shifted vertically, but their shapes and slopes are fixed by the temporal-integration data. Thus, the average amount of loudness summation, defined as level difference between equally loud tones and noises, may be adjusted to agree with the data, but the predicted level dependence of loudness summation is determined by the temporal-integration data. As shown in Fig. 6, the amount of loudness summation predicted from the loudness functions shown in Fig. 5 agrees closely with Pollack's (1951) data for loudness balance between white noise and 1-kHz tones. It is also in reasonable agreement with the amount of loudness summation derived from Scharf's (1978) loudness functions for 1-kHz tones and white noises. Considering that Florentine *et al.*'s (1996) listeners never heard tones together with the noises and that no attempt was made to have them relate loudness of one to the loudness of the other, we find this prediction of loudness summation quite remarkable. Moreover, because the loudness ratio between the 5- and the 200-ms stimuli is the same for tones and noises, the equal-loudness-ratio model also predicts that the amount of loudness summation should be independent of duration, which agrees with the data (Port, 1963; Zwicker, 1965). Altogether, these predictions offer strong support of the equal-loudness-ratio model. It should be noted, however, that little evidence is available to support the relatively large exponent of about 0.43 obtained for the 1-kHz tones at high levels on the basis of our temporal-integration data.

Finally, the equal-loudness-ratio model should also be able to predict Florentine *et al.*'s (1996) data for 30-ms stimuli. The loudness functions in Fig. 5 are determined only from the data for the 5- and 200-ms stimuli, but if the equal-loudness-ratio model is correct, the loudness functions for the 30-ms stimuli should be parallel to those for the other durations and the loudness ratios should be the same for tones and noises. Thus, only one free parameter is available to predict Florentine *et al.*'s (1996) data for the amount of temporal integration between 5- and 30-ms stimuli and between 30- and 200-ms stimuli for both tones and noises. In other words, with only one free parameter the model should predict four functions based on a total of 76 data points.

As shown in Fig. 7, the predicted functions are in good agreement with the polynomials used by Florentine *et al.* (1996) to summarize their data. The difference between the data and the predictions is always less than 2 dB, which is well within the variability of the data. The tendency for the predicted amounts of temporal integration to exceed that measured reflects the minor deviations from perfect transitivity observed in our previous data. The sum of level differences between 5- and 30-ms stimuli and between 30- and 200-ms stimuli tended to be a few dB less than that between 5- and 200-ms stimuli. In contrast, the predictions maintain

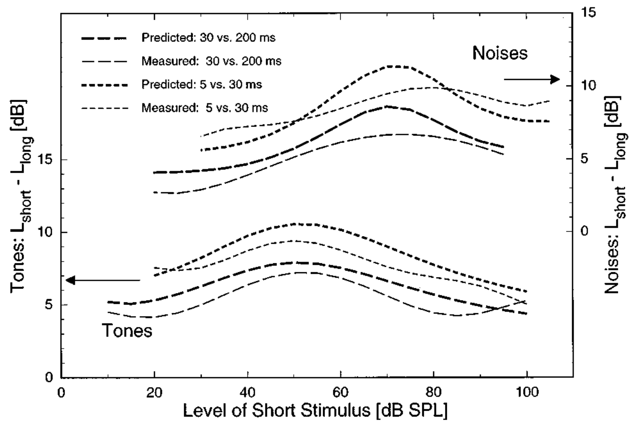

FIG. 7. Data and predictions for Florentine *et al.*'s (1996) 30-ms stimuli. The level difference between equally loud 5- and 30-ms stimuli (short-dashed lines) and between equally loud 30- and 200-ms stimuli (long-dashed lines) is plotted as a function of level of the shorter stimulus in the pair. The thick lines show predictions by the equal-loudness-ratio model and the thin lines show Florentine *et al.*'s (1996) summary of their data. The functions at the bottom are for tones and refer to the left ordinate; the functions at the top are for noises and refer to the right ordinate.

perfect transitivity. Because the level difference between 5- and 200-ms tones is fixed by the loudness functions shown in Fig. 5, the model was forced to exaggerate slightly the level differences between the 30-ms stimuli and the other stimuli. Overall, the forms of the predicted functions match those of the data very well. Thus, it appears that the loudness functions shown in Fig. 5 and the equal-loudness-ratio model can account for the effect of level on the amount of temporal integration observed between 5- and 30-ms stimuli and between 30- and 200-ms stimuli.

D. Loudness discrimination compared to level discrimination

It is often assumed that the just-noticeable difference in sound level, ΔL_{DL}, represents a difference limen for loudness (e.g., Harris, 1963; Hellman and Hellman, 1990; Neely and Allen, 1996). To examine this assumption, previous data for level discrimination as a function of duration may be compared with the present data for the just-noticeable loudness-level difference for variable stimuli of different durations. When listeners are asked to detect a small difference in level between two equal-duration tones, the ΔL_{DL} typically decreases by a factor of about 2 for each tenfold increase in duration (Florentine, 1986; Buus and Florentine, 1992). On the other hand, when listeners have to pick the louder of two tones with unequal duration, the jnds for loudness level appear to be independent of duration as indicated by the present data and the model. This result differs markedly from the expectation that the loudness-level jnds should be at least a factor of 4 smaller for 250-ms variable tones than for 2-ms variable tones, if level discrimination and loudness discrimination were governed by the same processes. This indicates that the ΔL_{DL}s measured in level-discrimination experiments in which the stimuli differ only in level are unlikely to correspond to a just-noticeable difference in loudness.[4]

The difference between level discrimination and loudness discrimination may be understood by considering the information available for each kind of discrimination. The Multiband Excitation-Pattern Model for Level Discrimination (Florentine and Buus, 1981) indicates that the excitation-level difference in each critical band provides highly reliable information for discrimination between two tones that differ only in level. The excitation-level difference in each critical band presumably can be observed for the duration of the stimuli, such that the number of observations increases in proportion to the stimulus duration. In fact, the increased number of observations accounts for the decrease in the jnds if the effects of auditory-nerve adaptation are taken into account (Buus and Florentine, 1992).

On the other hand, when the stimuli differ in several aspects, as in the present experiment, excitation-level differences in any one critical band are not reliable indicators of a change in level of the variable stimulus. Then, the listeners' judgments are likely to reflect an overall quality such as loudness. Indeed, our listeners were instructed to judge the loudness of the stimuli and probably their judgments were governed by loudness because the data indicate a close relation between the jnds and the loudness function. Difference limens based on judgments of loudness may be expected to be independent of the stimulus duration. Because the loudness is thought to represent the maximal output of an integrator (e.g., Zwicker, 1977; Zwicker and Fastl, 1990), the number of observations does not increase with stimulus durations—at least not when the stimuli are briefer than the relatively long time constant of the integrator [45 ms (Zwicker and Fastl, 1990) to 120 ms (Boone, 1973)]. Thus, the duration independence of the difference limens in the present experiment may reflect that loudness represents just one observation for both long and short stimuli. These considerations indicate that the difference between level jnds and loudness jnds is more than semantic.

IV. SUMMARY

The present study applied a new loudness-balance procedure to investigate temporal integration of loudness for 5-kHz tone bursts with durations ranging from 2 to 250 ms and levels ranging between 2 and 60 dB SL. The new procedure used ten interleaved adaptive tracks to obtain concurrent loudness balances at several levels with both the short and the long tone varied. It was designed to increase the consistency of the listeners' judgments by discouraging comparisons with previous trials. The data were interpreted in terms of an equal-loudness-ratio model, which assumes that the loudness ratio between two tones of different durations is independent of level and that the just-noticeable difference in loudness is proportional to the square root of loudness. The main conclusions are as follows:

(1) The new procedure appears successful as indicated by a high degree of internal consistency in the data. The transitivity of the loudness judgments appears nearly perfect.
(2) The amount of temporal integration, defined as the level difference between equally loud short and long tones, depends markedly on level. With the very brief tones

used in the present study, it varied more than 25 dB—from about 14 dB near threshold to over 40 dB when the 250-ms tone is about 22 dB SPL and the 2-ms tone is about 63 dB SPL. For all duration pairs, the amount of temporal integration is about three times larger at moderate levels than at low and high levels.

(3) The just-noticeable difference in loudness level also varies with level and is largest at moderate levels.

(4) Both effects of level can be explained by the equal-loudness-ratio model by assuming that the loudness function [log(loudness) as a function of SPL] is flatter at moderate levels than at low and high levels.

(5) The data and the model indicate that the just-noticeable difference for loudness level is independent of stimulus duration. This finding contrasts with previous data for level discrimination and indicates that level discrimination between stimuli that differ only in level is not the same as loudness discrimination between stimuli that differ along multiple dimensions.

(6) The equal-loudness-ratio model can predict the effect of level on loudness summation from previous data for temporal integration of loudness for 1-kHz tones and white noises. This finding offers strong support for the model because the measurements did not attempt to relate the loudness of white noise to that for tones.

(7) The equal-loudness-ratio model applied to data obtained with 5- and 200-ms stimuli also produces good predictions of the amount of temporal integration obtained with 30-ms stimuli.

ACKNOWLEDGMENTS

Part of this research was carried out at the Technical University of Denmark while the first and second authors were on sabbatical and leave of absence, respectively, from Northeastern University. It was supported by NIH-NIDCD R01 DC02241 and the Technical University of Denmark. The authors thank Rhona Hellman, Hannes Müsch, Bertram Scharf, Fan-Gang Zeng, and an anonymous reviewer for helpful comments on an earlier version of this paper.

[1]The reader should note that the term ''loudness discrimination'' is used to describe measurements of difference limens in tasks requiring the listeners to judge the loudness of stimuli that differ in more than one dimension (e.g., duration and level). The term ''level discrimination'' is used to describe measurements of difference limens for stimuli that differ only in level. Likewise, jnds for loudness level are used to indicate the difference limens (in dB) obtained for loudness discrimination, whereas jnds for sound level, $\Delta L_{\rm DL}$s, are used to indicate the just-audible difference between two sounds that differ only in level, as measured in traditional level- (or intensity-) discrimination experiments.

[2]The just-noticeable difference in loudness, $\Delta N_{\rm DL}$, often has been suggested to follow Weber's law such that $\Delta N_{\rm DL}/N$ (where N is the loudness in sones) is constant (e.g., Zwicker and Feldtkeller, 1967). Using this assumption in the model results in a fit that is only marginally poorer than the fit obtained when the just-noticeable difference in loudness is assumed to be proportional to the square root of the loudness.

[3]The Weber fraction for loudness, $\Delta N_{\rm DL}/N$, could conceivably depend on duration. However, the data show no clear effect of duration on the just-noticeable loudness-level differences. Moreover, models in which ∂ depended on duration showed no systematic effect of duration on ∂ and they provided only a marginally better fit to the data than the model in which ∂ is independent of duration. Therefore, ∂ was made independent of duration in the final model.

[4]It should be noted that Garner (1949) measured difference limens for loudness level by loudness balances between tones of different durations. Contrary to the present data, Garner's data indicate that the difference limen, calculated from the variability of loudness matches by the method of adjustment, increased as the duration of the short tone decreased. However, Garner's listeners varied only the long tone, which always was 500 ms in duration. Therefore, Garner's effect of duration may be confounded by effects of the level of the variable tone, which makes his data difficult to interpret. Moreover, when the two tones being compared had the same duration and the pause between them was 500 ms, only a small effect of duration on the differences limens was obtained—contrary to the effect of duration obtained in modern measurements of level discrimination (Florentine, 1986; Buus and Florentine, 1992). Thus, it appears that Garner's measurements with the method of adjustment are not comparable to measurements with modern forced-choice procedures.

Boone, M. M. (**1973**). ''Loudness measurements on pure tone and broad band impulsive sounds,''Acustica **29**, 198–204.

Buus, S., and Florentine, M. (**1992**). ''Possible relation of auditory-nerve adaptation to slow improvement in level discrimination with increasing duration,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Démany, and K. Horner (Pergamon, New York), pp. 279–288.

Cooper, N. P., and Yates, G. K. (**1994**). ''Nonlinear input-output functions derived from the responses of guinea-pig cochlear nerve fibers: Variations with characteristic frequency,'' Hear. Res. **78**, 221–234.

Durlach, N. I., and Braida, L. D. (**1969**). ''Intensity perception. I. Preliminary theory of intensity resolution,''J. Acoust. Soc. Am. **46**, 372–383.

Ekman, G., Berglund, B., and Berglund, U. (**1966**). ''Loudness as a function of the duration of auditory stimulation,'' Scand. J. Psychol. **7**, 201–208.

Fletcher, H., and Munson, W. A. (**1933**). ''Loudness, its definition, measurement and calculation,''J. Acoust. Soc. Am. **5**, 82–108.

Florentine, M. (**1986**). ''Level discrimination as a function of duration,''J. Acoust. Soc. Am. **79**, 792–798.

Florentine, M., and Buus, S. (**1981**). ''An excitation-pattern model for intensity discrimination,''J. Acoust. Soc. Am. **70**, 1646–1654.

Florentine, M., Buus, S., and Poulsen, T. (**1996**). ''Temporal integration of loudness as a function of level,''J. Acoust. Soc. Am. **99**, 1633–1644.

Garner, W. R. (**1949**). ''The loudness and loudness matching of short tones,'' J. Acoust. Soc. Am. **21**, 398–403.

Harris, J. D. (**1963**). ''Loudness discrimination,'' J. Speech Hear. Dis. Monogr. **11**.

Hellman, W., and Hellman, R. P. (**1990**). ''Intensity discrimination as the driving force for loudness. Application to pure tones in quiet,'' J. Acoust. Soc. Am. **87**, 1255–1271.

ISO 226 (**1987**). *Acoustics—Normal equal-loudness level contours* (International Organization for Standardization, Geneva).

ISO 389 (**1991**). *Acoustics—Standard reference zero for the calibration of pure-tone air conduction audiometers* (International Organization for Standardization, Geneva).

ISO TR 4869-3 (**1989**). *Acoustics—hearing protectors—part 3: Simplified method for the measurement of insertion loss of ear-muff type protectors for quality inspection purposes* (International Organization for Standardization, Geneva).

Jesteadt, W. (**1980**). ''An adaptive procedure for subjective judgments,'' Percept. Psychophys. **28**, 85–88.

Levitt, H. (**1971**). ''Transformed up–down procedures in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

Munson, W. A. (**1947**). ''The growth of auditory sensation,'' J. Acoust. Soc. Am. **19**, 584–591.

Neely, S. T., and Allen, J. B. (**1996**). ''Relationship between the rate of growth of loudness and the intensity DL,'' in *Modeling Sensorineural Hearing Loss*, edited by W. Jesteadt (Erlbaum, Hillsdale, NJ), in press.

Pedersen, O. J., Lyregaard, P. E., and Poulsen, T. (**1977**). ''The round robin test on impulsive noise,'' Report No. 22 (The Acoustics Laboratory, Technical University of Denmark, Lyngby).

Pollack, I. (**1951**). ''On the measurement of the loudness of white noise,'' J. Acoust. Soc. Am. **23**, 654–657.

Port, E. (**1963**). ''Über die Lautstärke einzelner kurzer Schallimpulse,'' Acustica **13**, 212–223.

Poulsen, T. (**1981**). ''Loudness of tone pulses in a free field,'' J. Acoust. Soc. Am. **69**, 1786–1790.

Scharf, B. (**1970**). ''Critical bands,'' in *Foundations of Modern Auditory Theory*, edited by J. V. Tobias (Academic, New York), Vol. I, pp. 157–202.

Scharf, B. (**1978**). ''Loudness,'' in *Handbook of Perception: Vol. 4. Hearing*, edited by E. C. Carterette and M. P. Friedman (Academic, New York), pp. 187–242.

Schlauch, R. S., and Wier, C. C. (**1987**). ''A method for relating loudness-matching and intensity-discrimination data,'' J. Speech Hear. Res. **30**, 13–20.

Stephens, S. D. G. (**1973**). ''Auditory temporal integration as a function of intensity,'' J. Sound Vib. **30**, 109–126.

Stevens, J. C., and Hall, J. W. (**1966**). ''Brightness and loudness as a function of stimulus duration,''Percept. Psychophys. **1**, 319–327.

Stevens, S. S. (**1956**). ''Calculation of the loudness of complex noise,'' J. Acoust. Soc. Am. **28**, 807–832.

Stevens, S. S., and Davis, H. (**1938**). *Hearing. Its Psychology and Physiology* (Wiley, New York; reprinted by AIP, New York, 1983).

Stevens, S. S., and Greenbaum, H. B. (**1966**). ''Regression effect in psychophysical judgment,'' Percept. Psychophys. **1**, 439–446.

Villchur, E. (**1969**). ''Free-field calibration of earphones,'' J. Acoust. Soc. Am. **46**, 1527–1534.

Yates, G. K. (**1990**). ''The basilar membrane input-output function,'' in *The Mechanics and Biophysics of Hearing*, edited by P. Dallos, C. D. Geisler, J. M. Matthews, M. A. Ruggero, and C. R. Steele (Springer-Verlag, New York), pp. 106–113.

Zeng, F.-G. (**1994**). ''Loudness growth in forward masking: Relation to intensity discrimination,'' J. Acoust. Soc. Am. **96**, 2127–2132.

Zwicker, E. (**1961**). ''Subdivision of the audible frequency range into critical bands (Frequenzgruppen),'' J. Acoust. Soc. Am. **33**, 248.

Zwicker, E. (**1965**). ''Temporal effects in simultaneous masking and loudness,'' J. Acoust. Soc. Am. **38**, 132–141.

Zwicker, E. (**1977**). ''Procedure for calculating loudness of temporally variable sounds,'' J. Acoust. Soc. Am. **62**, 675–682.

Zwicker, E., and Fastl, H. (**1990**). *Psychoacoustics: Facts and Models* (Springer-Verlag, New York).

Zwicker, E., and Feldtkeller, R. (**1967**). *Das Ohr als Nachrichtenempfänger* (Hirzel, Stuttgart).

Zwislocki, J. J. (**1965**). ''Analysis of some auditory characteristics,'' in *Handbook of Mathematical Psychology*, edited by R. D. Luce, R. R. Bush, and E. Galanter (Wiley, New York), pp. 1–97.

Zwislocki, J. J. (**1969**). ''Temporal summation of loudness: An analysis,'' J. Acoust. Soc. Am. **46**, 431–441.