

Lippincott
Williams & Wilkins**EAR and HEARING**
The Official Journal of the American AuditoryHome Search Current Issue **Archive** Publish Ahead of Print**October 1997, 18:5** > The Contour Test of Loudness Perception.< **Previous****ARTICLE LINKS:**[Abstract](#) | [References \(45\)](#) | [View full-size inline images](#)

Ear & Hearing:Volume 18(5)October 1997pp 388-400

The Contour Test of Loudness Perception
[Articles]

Cox, Robyn M.; Alexander, Genevieve C.; Taylor, Izel M.; Gray, Ginger A.

University of Memphis (R.M.C., G.G.), and Department of Veterans Affairs Medical Center (R.M.C., G.C.A., I.M.T.),
Memphis, Tennessee.

Address for correspondence: Robyn M. Cox, Speech and Hearing Center, 807 Jefferson Ave., Memphis, TN 38105.

Received September 9, 1996; accepted March 4, 1997

Abstract [TOP](#)

Objective: This article presents the underlying rationale, normative data, and reliability data for a test of loudness perception (the Contour Test) that was devised for use in clinical hearing aid fitting. The Contour Test yields data describing the sound level required for each of seven categories of loudness ranging from very soft to uncomfortably loud.

Design: Two experiments are described. Experiment 1 yielded norms for the test. The subjects were 23 male and 22 female normal-hearing listeners. Test stimuli included warble tones at six frequencies and broad band speech. Experiment 2 assessed the reliability of the test results. Ten hearing-impaired listeners responded to the test at two frequencies on two occasions separated by several days. Both experiments also evaluated the effect of using different stimulus increment sizes on the measured levels of loudness categories.

Results: Based on the data from experiment 1, norms for each category of each stimulus are reported in terms of mean level and typical between-subject variation in responses. Data are provided in HA-1 2 cm³ coupler levels as well as in hearing levels (dB HL). The shape of the loudness growth function for warble tones was somewhat different from that for speech. When data were expressed in HL, there were no differences in mean loudness category levels across warble tone test frequencies. Thus, test frequencies were combined and equations were generated to describe the upper and lower limits of typical normal performance for warble tone stimuli. These equations can be used to construct a template for clinical comparison of normative values to patient loudness growth curves. Experiment 2 provided information about the test-retest variability of data yielded by the Contour Test. Reliability appears to be similar to that of the few other category scaling tests described in the literature. Most test-retest differences were 6 dB or less. Although a moderate variation in test increment size did not significantly affect the loudness category levels for young normal-hearing listeners, levels corresponding to loudness categories were significantly higher when larger increments were used with elderly hearing-impaired listeners.

Conclusions: Evidence from this and other research indicates

Article Outline

- [Abstract](#)
- [Design Goals for the Contour Test](#)
 - [Psychophysical Parameters and Rationales](#)
- [Experiment 1: Contour Norms](#)
 - [Subjects](#)
 - [Procedure](#)
 - [Results and Discussion](#)
- [Experiment 2: Test-Retest Reliability](#)
 - [Subjects](#)
 - [Procedure](#)
 - [Results and Discussion](#)
- [General Discussion](#)
 - [Loudness Tests are not Commutable](#)
 - [Accounting for Loudness Summation Across Bandwidth](#)
- [Conclusion](#)
- [Acknowledgments:](#)
- [References](#)
 - [Reference Notes](#)

Figures/Tables

- [Figure 1](#)
- [Figure 2](#)
- [Table 1](#)
- [Table 2](#)
- [Figure 3](#)
- [Equation 1](#)
- [Equation 2](#)
- [Figure 4](#)
- [Figure 5](#)
- [Table 3](#)
- [Figure 6](#)

that standardized measurement of loudness perception is an achievable goal for clinical practice. The Contour Test appears to offer a viable approach to clinical measurement of loudness perception: It has good patient acceptance and combines fairly rapid administration with acceptable reliability. Details of test procedures and scoring sheets for manual administration can be down-loaded from the Internet at www.ausp.memphis.edu/harl. However, it is important to keep in mind that the application of loudness perception data for narrowband stimuli (such as warble tones) to hearing aid prescription is complicated by the need to account for the effects of loudness summation across bandwidth. There is a need for additional research to establish an empirical link between clinically measured loudness perception and optimal amplification characteristics.

- [Figure 7](#)
- [Table 4](#)
- [Figure 8](#)
- [Figure 9](#)
- [Table 5](#)

The appearance of a new generation of full dynamic range compression (FDRC) hearing aids has rekindled interest in the problems of quantification of loudness perception for individual patients in a clinical setting. Although consensus has not been reached about the characteristics of FDRC processing that will produce optimal performance, it is clear that this type of hearing aid has the potential to address the abnormalities in loudness growth that are the signature of a hearing loss of cochlear origin. It seems reasonable to hope that with FDRC processing, sounds that are soft, comfortable, and loud for a normal-hearing listener can be made soft, comfortable, and loud, respectively, for the hearing aid wearer without the need for constant volume control adjustments. On the other hand, it must be admitted that the mere restoration of relatively normal loudness experiences may not produce optimal speech recognition performance. There is continuing discussion about the specific applications of loudness perception measures to hearing aid fitting (e.g., [Byrne, 1996](#); [Grey & Dyrland, 1996](#); [Ricketts, 1996](#)). Nevertheless, many practitioners feel that attention to loudness perception issues is an essential first step in utilization of the new nonlinear devices for a large number of hearing aid candidates.

The lack of a widely accepted method for clinical quantification of suprathreshold loudness perceptions is one factor that might limit the acceptance and employment of FDRC hearing aids in the field. In an attempt to fill this void, investigators have put forth a variety of procedures. This article describes one such procedure, the Contour Test. This test has been developed specifically for application to hearing aid fitting in a clinical setting. We present the design goals, normative data, test-retest reliability, and some potential applications.

Design Goals for the Contour Test [TOP](#)

There has been much debate about the "correct" approach to loudness measurement for application to hearing aid fitting. The position taken here is that there is no single most accurate psychophysical method. There is ample evidence that an absolute, internal representation of loudness does not exist. Instead, the results of any particular testing procedure are substantially influenced by the details of the measurement methods (e.g., see [Filion & Margolis, 1992](#) and [Mellers, 1983](#)). Thus, efforts should be directed towards the design of a scientifically defensible procedure that is, at the same time, clinically practicable. The results of this procedure can then be related to hearing aid fitting using empirical methods. The most important consideration is the establishment of a standardized protocol for loudness perception measurement that is employed in a systematic manner across clinical sites.

This problem has many points of similarity with the effort to develop standardized methods for measurement of sensitivity thresholds, which culminated with an American National Standard ([American National Standards Institute, 1978](#)) recommending universal adoption of a variant of the procedure described by [Carhart and Jerger \(1959\)](#). The use of this procedure for threshold measurement has now become so second nature in audiometry that clinicians might have forgotten that it has no intrinsic validity. The fact that thresholds measured using different procedures could be different to some extent is of no importance because thresholds measured using the standard procedure have been related to diagnosis and treatment using empirical methods. The establishment and acceptance of an analogous procedure for loudness perception measurement is a worthy goal. The details of the Contour Test described in this paper were developed with these considerations in mind.

Psychophysical Parameters and Rationales [TOP](#)

Loudness Growth Function • Because FDRC hearing aids have the potential to apply nonlinear processing to the entire range of input levels experienced by the hearing aid wearer, the loudness test was designed to develop a function that describes the growth of loudness perception as input levels increase from near threshold to uncomfortably loud. It is then possible to use the entire function, or any specific points on the function, to apply to the hearing aid prescription. At least two approaches to determining loudness growth functions appear to meet the requirement for clinical feasibility. They are category ratings of loudness comfort (e.g., [Allen, Hall, & Jeng, 1990](#); [Pascoe, 1986](#)) and restricted magnitude estimation of loudness (e.g., [Cornelisse & Jamieson, Reference Note 1](#); [Elberling & Nielsen, 1993](#); [Geller & Margolis, 1984](#)). Although both approaches have arguable merits, category ratings of loudness comfort were chosen for the Contour Test on the grounds that this method might be somewhat easier, and therefore faster, for the hearing-impaired patient. Seven categories of loudness were selected from the nine proposed by [Hawkins, Walden, Montgomery, and Prosek \(1987\)](#). [Figure 1](#) is a facsimile of the 8.5 × 11 card displaying the categories that the patient is given to refer to during the test.

LOUDNESS CATEGORIES

7. Uncomfortably Loud
 6. Loud, but O.K.
 5. Comfortable, but Slightly Loud
 4. Comfortable
 3. Comfortable, but Slightly Soft
 2. Soft
 1. Very Soft
-

Figure 1. Loudness category numbers and descriptors shown to the subject during testing.

Instructions • The instructions for any loudness test are a critical factor in determining the results (e.g., [Beattie, Svihovic, Carmen, & Kunkel, 1980](#)). This is so important that we recommend that the instructions be written down and adhered to very carefully. On the other hand, giving a patient a page of written instructions and saying "read these" often does not result in a good test, probably because the patient is not able to absorb the instructions under the pressure of the moment. An approach that we have used with success is to reproduce the instructions in large type and read them over with the patient in a low-key tutorial fashion, using gestures and repetition as necessary. The instructions for the Contour Test are as follows:

"The purpose of this test is to find your judgments of the loudness of different sounds. You will hear sounds that increase and decrease in volume. You must make a judgment about how loud these sounds are. Pretend you are listening to the radio at that volume. How loud would it be? After each sound, tell me which of these categories best describes the loudness. Keep in mind that an uncomfortably loud sound is louder than you would ever choose on your radio no matter what mood you are in." *

After these instructions have been reviewed with the patient, it is important to verbally add that he or she should use any category that seems appropriate at any time and that it is permissible to repeat a category or to skip categories. Further elaboration of the instructions should be kept to a minimum.

Stimulus • To provide a basis for prescribing hearing aid characteristics as a function of frequency, the stimulus used for loudness testing must be restricted in bandwidth. For the most part, stimuli more narrow than a normal critical band have been used for this task, with the most common choices being pure tones, warble tones, noise bands of about 1/3 octave, and 1/3 octave bands of filtered speech. Several studies have shown that pure tones, warble tones, and 1/3 octave noise bands produce the same results in loudness tests, whereas results for 1/3 octave speech bands have been less consistent (e.g., [Byrne, 1986](#); [Cox, 1989](#); [Hawkins, 1980](#); [Ricketts & Bentler, 1996](#)). Warble tones were chosen as the preferred stimulus for the Contour Test because they sample a somewhat wider bandwidth than pure tones while avoiding the pitfalls associated with the more gradual filter slope often found in noise bands produced by audiometers (e.g., see [Orchik & Mosher, 1975](#)). However, under most conditions pulsed pure tones or 1/3 octave noise bands could be substituted for the warble tones without affecting the outcome.

Stimulus duration also affects loudness judgments ([Cox, 1989](#)) and the effects of duration have been shown to differ with stimulus type ([Florentine, Buus, & Poulsen, 1996](#)). Thus, although there may be no obviously optimal stimulus duration for use in loudness tests, it is clearly important to control this variable. In the Contour Test the warble tones are presented in groups of four 200 msec pulses with a 50% duty cycle. This is similar to the automatically pulsed stimulus found on many audiometers.

Level Sequence • There is no doubt that the apparent loudness of a stimulus on a particular trial is influenced by the levels of stimuli that have been presented in previous trials. In particular, several investigations have shown that the loudness judgment allocated to a stimulus on a particular trial tends to be "attracted by," or similar to, the loudness of the stimulus on the preceding trial (e.g., [Jesteadt, Luce, & Green, 1977](#); [Ward & Lockhead, 1970](#)). To explore sequencing effects, several investigators have contrasted the results obtained with a strictly ascending and a strictly descending series of stimuli (e.g., [Ventry & Johnson, 1978](#); [Woods, Ventry, & Gatling, 1973](#)). These studies demonstrate that, when differences are seen, a given loudness category will always occur at a lower level in an ascending series than in a descending series. In addition, [Jenstad, Cornelisse, and Seewald \(Reference Note 6\)](#) have demonstrated that the sequencing of stimuli has a systematic effect on the slope of the derived loudness function for an individual. However, some subjects appear to be minimally affected by the sequencing context (e.g., see [Cornelisse & Jameson, Reference Note 1](#); [Knier & Bentler, Reference Note 8](#)). Some investigators have elected to use randomized level sequencing in loudness testing (e.g., [Allen, Hall, & Jeng, 1990](#)). This is probably intended to reduce sequence effects if data are averaged across a large number of trials with randomly changing levels. However, [Hellbruck, Thomamuller, and Zeitler \(1995\)](#) have shown that sequence effects are still clearly in evidence in the data from tests with randomized sequencing employing 24 presentations (12 per frequency).

The effects of stimuli sequence are apparently ubiquitous and unavoidable in loudness judgments and there is no obviously "correct" sequencing mode. It is, therefore, of considerable importance to specify the sequencing of stimuli for a clinical loudness test and to adhere to it in all applications of the test. In many cases, different sequencing produces a different result. In the Contour Test, the stimulus sequencing is strictly ascending, that is, a run begins at a level just above threshold and the level of each successive stimulus is raised, with the patient furnishing a

loudness category judgment for each stimulus, until a judgment of uncomfortably loud is given. Ascending sequencing was chosen because both patients and clinicians are comfortable with it and because it is readily implemented manually as well as under computer control.

Stimulus Increment • Clinical loudness tests have employed a wide range of level increments and at the outset it was not clear whether the choice of increment size is a critical variable. To explore this issue in the development phase, a pilot study was performed using two normal-hearing listeners and two versions of the Contour Test. In version 1, stimulus increments were varied randomly among 2, 3, 4, and 5 dB. In version 2, stimulus increments were varied randomly among 3, 5, 7, and 9 dB. **Figure 2** depicts the results for the very soft (vs); comfortable, but slightly soft (css); comfortable, but slightly loud (csl); and uncomfortably loud (ucl) contours for the subject showing the smallest effect of increment size. There is a clear trend for the level allocated to a given loudness category to be higher in the test with larger increments. Results were similar for both subjects, suggesting that larger increments between stimuli result in contours at higher levels. We concluded that it is probably important to control increment size and explored this variable further in experiments 1 and 2. Based on the results of these experiments, increment sizes for the final version of the Contour Test were chosen to provide resolution of seven contours within the available dynamic range as follows: if threshold is less than 50 dB HL, increment size is 5 dB; if threshold is greater than, or equal to, 50 dB HL, increment size is randomly chosen among 2, 3, 4, and 5 dB for computer administration or fixed at 2 or 2.5 dB (depending on audiometer) for manual administration.

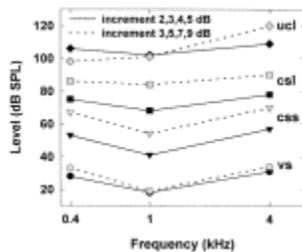


Figure 2. Contour data obtained for one subject using two increment sizes. vs = very soft; css = comfortable, but slightly soft; csl = comfortable, but slightly loud; ucl = uncomfortably loud.

Experiment 1: Contour Norms [TOP](#)

Subjects [TOP](#)

Forty-five individuals participated. All were from outside the field of communication disorders and were naïve with respect to loudness perception testing. Ages ranged from 18 to 37 with a mean of 24 yr. There were 23 men and 22 women. All subjects reported having normal hearing. Sensitivity thresholds were determined using an ER-3A insert earphone coupled to the test ear with a compressible foam plug. The non-test ear was also plugged. Earphone output was calibrated in sound pressure level in an HA-1 2 cm³ coupler. Threshold data are given in [Table 1](#).

Frequency (Hz)	Mean Threshold (dB SPL)		SD (dB)	
	Male Subjects	Female Subjects	Male Subjects	Female Subjects
250	22.9	21.4	5.1	4.4
500	13.0	11.5	5.5	5.5
1000	8.1	7.1	4.9	4.4
2000	8.4	10.8	3.2	4.8
3000	8.8	8.4	4.0	4.3
4000	10.9	12.6	5.1	5.1

TABLE 1. Means and SDs of thresholds for male and female subjects in Experiment 1 at each test frequency. Sound pressure levels are referenced to an HA-1 coupler.

Procedure [TOP](#)

Tonal stimuli were 5% warble tones presented at six test frequencies (250, 500, 1000, 2000, 3000, and 4000 Hz). Duration and number of pulses were as described above (four 200 msec pulses with a 50% duty cycle). Order of frequency presentation was randomized.

In addition, for comparison purposes, testing was performed using connected speech. Each trial was composed of a 5.0 sec. sample of speech drawn from an ongoing passage of discourse from the Connected Speech Test ([Cox, Alexander, Gilmore & Pusakulich, 1989](#)). Thus, the precise speech sample was different for each trial. Half of the subjects responded to speech first and the other half heard the tones first. Twenty-three subjects were tested using a fixed 5 dB increment between stimuli and 22 were tested using increments randomly chosen among 2, 3, 4, and 5 dB.

Stimuli were delivered monaurally by the ER-3A insert earphone used for threshold testing. Both speech and warble tones were calibrated in terms of overall sound pressure level (the continuous speech was integrated over a 1-minute time interval including natural silent periods). A run began with a stimulus delivered one increment above threshold and continued in a strictly ascending fashion with the subject providing a loudness category for each stimulus. The run terminated when a judgment of uncomfortably loud was given. Before data collection began, each subject heard one run using the 1000 Hz warble tone for practice. Subsequently, each stimulus was tested using four consecutive runs. After the test, the value for each loudness category was computed as the median sound pressure level of responses using that category across the four runs.

Results and Discussion [TOP](#)

For each test frequency, mean loudness contour levels were subjected to repeated measures analysis of variance with two variables: increment strategy (fixed versus random) and loudness category (seven), to explore the effects of the fixed and random increments on the measured levels of loudness categories. No analysis produced a significant effect for increment, indicating that for normal-hearing subjects these two methods of setting increment size did not produce systematically different results (it should be noted that there was a nonsignificant trend for the mean fixed increment result to be about 2 dB higher than the mean random increment result). Based on this outcome, data for the two increment strategies were combined for further analysis.

At least two previous investigations have found that a sound of a particular level is given a higher loudness rating by women than by men ([Kiessling, Steffens, & Wagner, 1993](#); [Nielsen, Reference Note 9](#)). In both of these reports, the loudness function for normal-hearing men was separated from the corresponding function for normal-hearing women by about 6 dB. In the present data, the mean level across all loudness categories was 68.2 dB for women and 71.5 dB for men. Thus, the trend in the data was consistent with the studies cited above. To examine the significance of the male-female difference, an analysis of variance was performed with one between-groups variable (gender) and two within-groups variables (frequency and loudness category). The main effect of gender was not statistically significant ($p > 0.1$).

[Table 2](#) gives group mean loudness category levels and standard deviations for each stimulus, expressed in sound pressure levels in an HA-1 coupler. Data in this format can be directly compared with hearing aid performance measured in an HA-1 coupler (e.g., see [Cox, 1995](#)). Within each stimulus, the results in [Table 2](#) show the expected increase in mean levels as loudness categories increased. Also note that the mean overall speech level corresponding to any particular loudness category was almost always lower than any of the warble tone levels in that category. The largest difference between speech and warble tones is seen in the "comfortable, but slightly soft" category where the corresponding overall speech level is 41.9 dB SPL and the corresponding warble tone levels are 12 to 25 dB higher than this. Perhaps the temporal and/or durational differences between speech and warble tone stimuli contributed to this outcome. However, it seems probable that most of the difference in the loudness of speech and warble tone stimuli reflects level-dependent loudness summation across bandwidth for the speech stimulus. The implications of this outcome are discussed in more detail below.

TABLE 2. Mean loudness category levels and SDs, in parentheses, for each stimulus pressure levels are referenced to an HA-1 coupler.

The standard deviations in [Table 2](#) reveal that the between-subject variability was fairly similar for a given loudness category across warble tone test frequencies, but increased monotonically as loudness progressed from soft to loud. Within each category, the standard deviation was smaller for speech than for the warble tones, indicating that loudness judgments for speech were more consistent across listeners than those for warble tones. However, the variability increased across categories for speech in a pattern similar to that shown for warble tones. The implication of this regular increase in variability across categories is that normal-hearing listeners are more alike in their judgments of which sounds are soft than they are in their judgments of which sounds are loud. However, the fact that the ascending test procedure is anchored at the lower end of the scale might also promote greater consistency across subjects in the low loudness categories. When investigators have employed a test procedure with randomly varied stimulus levels, some (e.g., [Hellbruck et al., 1995](#)) report results suggesting that between-subject variability is greater for higher loudness levels than for lower ones but others (e.g., [Elberling & Neilsen, 1993](#)) have not observed this pattern.

The loudness category levels for warble tone stimuli in [Table 2](#) can be displayed as a multi-line plot showing each individual category level as a function of frequency. This creates a set of "equal comfort contours." It is perhaps worth noting that it is not necessarily accurate to refer to these data as equal loudness contours because two sounds that are equally comfortable might not be equally loud. We have suggested several ways in which equal comfort contours might be useful in hearing aid fitting ([Cox, Goff, Martin, & McLoud, Reference Note 2](#); [Cox, Taylor, Gray, & Brainerd, Reference Note 3](#); [Cox & Alexander, Reference Note 4](#)). However, most of the applications of loudness perception data to hearing aid fitting have used loudness growth curves. These depict loudness responses at all categories for one test stimulus.

It can be instructive to compare a hearing-impaired individual's loudness growth curve with the loudness growth pattern exhibited by typical normal-hearing listeners. To facilitate this type of comparison, the loudness growth data for each warble tone test frequency were converted to Hearing Levels by subtracting the reference threshold sound pressure levels for HA-1 coupler for the ER-3A earphone ([American National Standards Institute, 1989](#)). The results for all six test frequencies are displayed on the same coordinates in [Figure 3](#). The open symbols depict the mean warble tone level corresponding to each category. Each frequency is plotted using a different unfilled symbol. However, as the figure reveals, when data were expressed in dB HL, the loudness growth patterns were so similar across frequencies that it is not possible to distinguish the different symbols. Thus, it seemed appropriate to combine the data across frequencies.

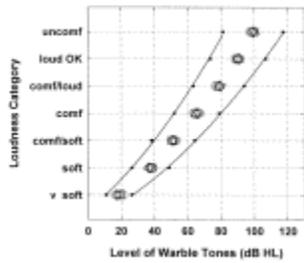


Figure 3. Open symbols depict the mean level (dB HL) for warble tone stimuli for each loudness category for normal-hearing listeners. Different symbols are used for each test frequency. Small filled circles show the upper and lower limits of the typical normal range of results. Solid lines have been fitted to the upper and lower limit values.

The small solid circles in [Figure 3](#) encompass the middle 80% of the data and, therefore, denote the upper and lower boundaries of the typical normal range of levels associated with each loudness category. They were determined as follows: 1) the mean warble tone level for each category was computed by averaging across frequencies, 2) variability within each category was determined by combining the standard deviations for the six test frequencies on an RMS basis, and 3) the upper and lower limits of the middle 80% of the data were then determined based on the assumption of a normal distribution of data in each category.

To evaluate results from a hearing-impaired listener, it may be convenient to plot them on a template that displays, for comparison, the range of sound levels typically associated with each loudness category by normal-hearing listeners. To facilitate construction of a suitable clinical template, equations were derived for drawing smooth curves through the upper and lower limits (small filled circles) shown in [Figure 3](#). For these equations, the loudness categories "very soft" to "uncomfortably loud" were denoted by integer values from 1 to 7, respectively ([Fig. 1](#) shows the numbers associated with each category). The upper limit values were fitted with a second order polynomial. Another second order polynomial was fitted to the lower limit values. Both polynomials produced a correlation of 0.999 with the underlying data and curves produced using them, shown as solid lines in [Figure 3](#), are an excellent description of the data points. For readers who wish to produce their own templates, the equations are: [Equation 1](#), [Equation 2](#)

$$\text{Loudness Category}^{\text{upper limit}} = 0.0003612(\text{HL}^2) + 0.0509(\text{HL}) + 0.423$$

Equation 1

$$\text{Loudness Category}^{\text{lower limit}} = 0.0002494(\text{HL}^2) + 0.0299(\text{HL}) + 0.014$$

Equation 2

The upper limit line can be constructed by assuming a series of values for HL and solving [Equation 1](#) to determine the loudness category value corresponding to each HL value. The lower limit line is constructed in the same manner using [Equation 2](#). For both lines, any HL values that result in loudness categories less than 1 or more than 7 should be discarded. Note that these functions should be interpreted as depicting the range of signal levels that are typically associated with specific loudness categories. It would not be appropriate to use these functions as confidence regions for loudness categories that are associated with specific signal levels.

[Figure 4](#) depicts loudness growth data for the speech stimulus. The large unfilled circles show group mean sound levels for each loudness category and the small filled circles indicate the upper and lower boundaries of the middle 80% of the data. Polynomials were fitted to the upper and lower limits in the same way as described for the data in [Figure 3](#) and these are shown as solid lines. It was necessary to use a third order polynomial to fit the underlying data as well as in [Figure 3](#). If we compare [Figures 3 and 4](#), two differences are noteworthy. First, it appears that the shape of the loudness growth function is different for speech and warble tones. Second, the smaller variability associated with judgments of speech, noted above, is evident in the closer spacing of the lines delimiting the range of typical data. These differences are more clearly displayed in [Figure 5](#), which compares the ranges of typical data (the solid lines from [Figs. 3 and 4](#)) for both types of stimuli.

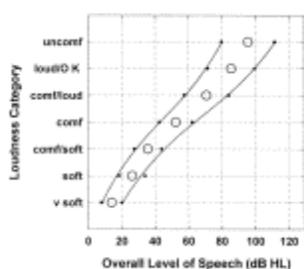


Figure 4. Open circles depict the mean level (dB HL) for broadband speech stimuli for each loudness category for normal-hearing listeners. Small filled circles show the upper and lower limits of the typical normal range of results. Solid lines have been fitted to the upper and lower limit values.

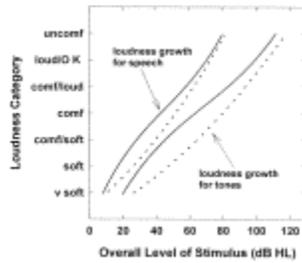


Figure 5. Solid lines depict the typical range of sound levels associated with each loudness category for speech stimuli. Dotted lines depict the corresponding range of sound levels for warble tone stimuli.

Consider the left solid line and the left dotted line in [Figure 5](#). These depict the corresponding lower limits of the normal ranges for speech stimuli and warble tone stimuli, respectively. Notice that the two lines almost coincide at the bottom and top, corresponding roughly to "very soft" = 10 dB HL and "uncomfortably loud" = 80 dB HL. However, between these points, the data for the speech stimulus (solid line) lies to the left of that for the warble tones (dotted line) with a maximum difference of about 11 dB for the category "comfortable but slightly soft." This outcome is consistent with the expected effects of loudness summation across bandwidth for the speech stimulus. [Zwicker, Flottorp, and Stevens \(1957\)](#) demonstrated that when the bandwidth of a sound of constant overall level is increased beyond a certain width (called the critical bandwidth for loudness summation), the loudness of the sound increases despite the fact that the overall level remains the same. They also noted that the magnitude of loudness summation across bandwidth is level-dependent (at least in normal-hearing listeners), reaching a maximum for sounds that are at medium levels and declining thereafter.

Experiment 2: Test-Retest Reliability [TOP](#)

If individual loudness perception data are to be used to prescribe and fit amplification, it is important for the clinician to be reasonably confident that the loudness category levels measured on the day of the evaluation would be repeated with acceptable accuracy on a subsequent test occasion. Further, although it is widely recognized that a loudness perception test for clinical application must satisfy requirements for administration time, it is perhaps less well recognized that there is a link between administration time and reliability. In general, a shorter test yields less reliable data. Thus, one challenge in the development of a clinical test of loudness perception is to strike an appropriate balance between administration time and data reliability. Experience with the Contour Test indicated that, once the clinician is fully conversant with the procedure, administration time is five minutes or less per test frequency. In this experiment, we investigated the reliability of the data. In addition, two test increment strategies were evaluated to determine whether small variations in test increment affected the outcome and whether either type of increment variation strategy resulted in a more reliable test.

Subjects [TOP](#)

Six men and four women with sensorineural hearing impairment served as subjects. Ages ranged from 70 to 85 with a mean of 77 yr. A variety of audiometric configurations were represented, ranging from mild-moderate-sloping to severe-flat.

Procedure [TOP](#)

Each subject was tested twice. Test-retest intervals ranged from three to fifteen days with a mean of 8 days. During each test session, the subject responded to two versions of the Contour Test: one version employed a fixed 5 dB stimulus increment and the other version used stimulus increments randomly chosen from 2, 3, 4, and 5 dB. Four frequencies were tested: 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz. Testing was unilateral. Administration details and instrumentation were the same as described for Experiment 1.

Results and Discussion [TOP](#)

Test-retest reliability was assessed in terms of the variability in stimulus levels judged to correspond to each loudness category. Data for each of the seven loudness categories were analyzed using repeated measures analysis of variance with three variables: test frequency (four), increment strategy (fixed or random), and test session (two). All analyses produced a significant main effect for test frequency ($df = 3, 27; p < 0.02$). However, this simply reflects the characteristic variations in sensitivity across frequency in these types of subjects and is of no interest for this investigation. Of much more interest, on the other hand, is the finding that the analysis for every category produced a significant effect for increment strategy ($df = 1, 9; p < 0.001$). This indicates that when test increments were randomly chosen from among 2, 3, 4, or 5 dB, the level for every loudness category was different from that obtained with a fixed 5 dB increment. This result obtained with elderly hearing-impaired listeners contrasts sharply with the result for young normal-hearing listeners obtained in experiment 1. [Table 3](#) gives the mean loudness category levels and the standard deviations obtained with each increment strategy and these data allow evaluation of the direction and extent of the differences. For most categories, the loudness level determined using the fixed 5 dB increment was 3 to 5 dB higher than the level determined for the same category using an increment that varied between 2 and 5 dB. Although different from the results of Experiment 1, this outcome is consistent with the finding of the pilot investigation that compared larger variable increments with smaller variable increments (illustrated in [Fig.](#)

2) in suggesting that a larger increment produces contours at higher levels.

Loudness Category	Fixed 5 dB Increment		Random 2-5 dB Increment	
	Mean (SD SPL)	SD (dB)	Mean (SD SPL)	SD (dB)
Very soft	68.8	12.5	67.3	12.9
Soft	77.2	11.7	74.3	12.0
Comfortable, but slightly soft	83.2	11.7	79.6	11.9
Comfortable	88.8	11.4	84.8	11.1
Comfortable, but slightly loud	94.5	10.8	90.3	11.1
Loud, but OK	97.4	10.6	92.4	11.4
Uncomfortably loud	102.3	10.4	96.9	11.8

TABLE 3. Means and SDs of loudness category levels obtained with fixed and random test increments from 10 hearing-impaired listeners. Sound pressure levels are referenced to an HA-1 coupler.

None of the analyses produced a significant effect for session, implying that there was not a systematic trend for loudness judgments to increase or decrease over time. Although this is important, it is not sufficient to guarantee adequate retest reliability for clinical purposes. To assess this, we must consider the distribution of test-retest differences on an individual basis.

Figures 6 and 7 address this issue. The upper panel of Figure 6 illustrates the test-retest differences for all 10 subjects and each loudness category for the 500 Hz test frequency and the fixed test increment. The lower panel gives the corresponding information for the 2000 Hz test frequency and the random test increment. Both panels show the same trends and are illustrative of the results for the other test frequencies. These results indicate that the individual test-retest differences for hearing-impaired listeners are typically smaller for the softer categories than the louder ones but, with a few exceptions, generally remain less than 10 dB for all categories. Figure 7 gives the complete distribution of test-retest differences for each increment type with data pooled across test frequencies and loudness categories. Both distributions reinforce the impression gained from Figure 6 and indicate that most test-retest differences do not exceed 6 dB. Differences exceeding 10 dB were seen only 1 to 2% of the time.

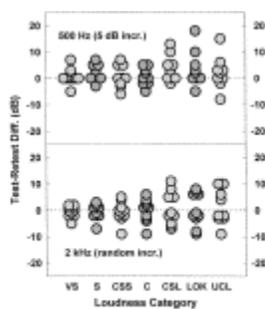


Figure 6. Test-retest differences for 10 subjects and each loudness category. The upper panel gives data for the 500 Hz warble tone and fixed 5 dB stimulus increment. The lower panel gives data for the 2000 Hz warble tone and the randomly varied stimulus increment.

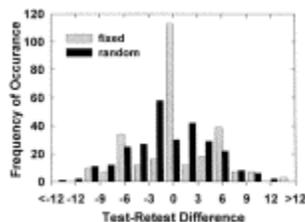


Figure 7. Distribution of test-retest differences for loudness test using fixed and randomly varied stimulus increments.

Some previous investigations have reported reliability in terms of the standard deviations of test-retest differences. For comparison with these data, standard deviations of test-retest differences were computed for each loudness category with data combined across test frequencies and increment strategies. The results are given in Table 4. These data also illustrate the trend noted in Figure 6 for test-retest variability to increase for higher loudness categories. The standard deviations range from about 2 dB to about 6 dB. These values are quite similar to the 3 dB to 7 dB reported by Robinson and Gatehouse (1996) for a somewhat more time-consuming category rating procedure.

Loudness Category	SD (dB)
Very soft	2.1
Soft	3.1
Comfortable, but slightly soft	4.0
Comfortable	4.4
Comfortable, but slightly loud	4.7
Loud, but OK	6.2
Uncomfortably loud	5.9

TABLE 4. SDs of test-retest differences for each loudness category. Data are combined across test frequencies and increment strategies.

General Discussion [TOP](#)

The results reported here support the following statements about the characteristics of data obtained with the Contour Test with warble tone stimuli:

1. The pattern and range of results for normal-hearing listeners is independent of test frequency when stimulus levels are calibrated in terms of hearing threshold level. Thus, a normative template can be generated, using the equations provided, which can be employed clinically at test frequencies ranging from 250 Hz to 4000 Hz. [Figures 8 and 9](#) illustrate the use of this kind of approach to assist in formulating an amplification strategy. Each figure depicts the Contour Test results for a different hearing-impaired individuals compared with the typical range of results seen for normal-hearing individuals. In [Figure 8](#), data for patient BM are similar for both test frequencies: the stimulus level required for a judgment of "very soft" is about 45 dB more than that needed by normal-hearing listeners. However, as stimulus level is increased above the very soft level, the difference between loudness judgments given by BM and those from listeners with normal hearing diminishes. Stimuli judged by BM to be "comfortable but slightly loud" are also given this loudness rating by normal-hearing listeners. All stimulus levels above this are also perceived with normal loudness by BM. Seeing this pattern (which is very typical), the clinician could reasonably conclude that BM will benefit from amplification for sounds that are normally perceived as soft but no gain is needed for sounds that are normally perceived as comfortable or louder.

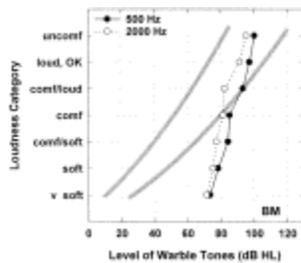


Figure 8. Contour data for Patient BM compared with the typical range of responses from normal-hearing listeners.

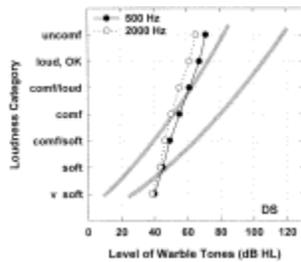


Figure 9. Contour data for Patient DS compared with the typical range of responses from normal-hearing listeners.

[Figure 9](#) illustrates a different pattern of results seen for patient DS. This patient presented with a complaint of difficulty with soft sounds and the Contour Test results confirmed that, for both test frequencies, sounds needed to be about 20 dB more than normal to produce a judgment of "very soft." However, the results also indicate that DS displayed a hypersensitivity to loud sounds. Responses in the higher loudness categories are seen to lie to the left of the normal range showing that levels that are ordinarily judged to be comfortably loud appeared uncomfortable to DS. This outcome implies that even unamplified everyday sounds might be too loud for this patient to endure comfortably and suggests a poor prognosis for hearing aid use.

2. Elderly hearing-impaired listeners' results were affected by small differences in stimulus increments despite the fact that these differences did not have a substantial effect on results for young listeners with normal hearing. We cannot determine from these data whether the effect of increment strategy (fixed versus variable; larger versus smaller, etc.) was related to hearing loss or age, or both. Nevertheless, this outcome reinforces the importance of using a standard procedure, including standard increment sizes, when applying loudness scaling to hearing aid fitting. Based on the results of Experiments 1 and 2, the standard increment chosen for the Contour Test is 5 dB when the hearing threshold is less than 50 dB HL and a smaller number when the threshold is poorer than 50 dB (the exact size of the smaller increment depends on the particular audiometer and whether the audiometer is controlled manually or by computer). The use of smaller increments for frequency regions with a smaller dynamic range is intended to promote greater resolution of the seven loudness contours than could be achieved with a 5 dB increment. The smaller increment can also be expected to result in loudness contours a few dB lower than would be obtained with a 5 dB increment. The consequence will be a somewhat more conservative hearing aid fitting (less gain and lower maximum output) at frequencies with small dynamic ranges. Further study will be required to evaluate the efficacy of this aspect of the procedure.

3. In the past, some clinicians have been hesitant to adopt loudness scaling as a routine part of their hearing aid fitting methods. One reason for this is the reputed poor reliability of loudness scaling data. However, several studies have shown that when a standardized method is consistently applied, the reliability of clinical loudness scaling is similar to that of clinical threshold testing ([Cox, 1989](#); [Sammeth, Birman, & Hecox, 1989](#)). Although there are relatively few published data that reflect the variability of stimulus levels that are judged to correspond to loudness scaling categories, the results obtained for the Contour Test appear to be similar to those reported for other tests. Some positive aspects of the Contour Test include its amenability to manual or computer-controlled administration and its relative time efficiency compared with other equally reliable procedures ([Robinson & Gatehouse, 1996](#)).

Overall, the Contour Test appears to be a reasonable approach to the quantification of loudness perceptions in a clinical setting.† However, the application of the results of loudness scaling to hearing aid fitting is not entirely straightforward. In particular, there are at least two matters that need careful consideration. First, different loudness testing procedures cannot be freely substituted for one another. Second, although practical considerations dictate that loudness perception tests must employ stimuli of restricted bandwidth, hearing aids in actual use are generally exposed to complex stimuli with greater (but unpredictable) bandwidths. Each of these matters is discussed more fully below.

Loudness Tests are not Commutable [TOP](#)

One of the most pressing applications for loudness scaling data is to assist in the prescription and fitting of hearing aids. This was the motivation for the development of the Contour Test. The VIOLA (Visual Input/Output Locator Algorithm) procedure (Cox, 1995) has been developed to utilize data obtained from the Contour Test for amplification prescription and hearing aid selection. Other loudness scaling procedures have been developed and integrated into different hearing aid prescription, selection, and fitting methods (e.g., Kiessling, Schubert, & Archut, 1996; Ricketts & Van Vliet, 1996). Because different loudness measurement methods may have superficial similarities, it is sometimes tempting to substitute loudness test data from one method into the prescriptive procedure for another method. However, this temptation should be resisted unless the loudness measurement procedures have been demonstrated to produce similar results. It is important to keep in mind that categories of loudness such as "comfortable," "very soft," etc., are somewhat fluid in the minds of most listeners. Thus, the levels corresponding to various categories that are obtained with any individual listener will be influenced by the details of the psychophysical procedure used in the measurement process. This means that it is not necessarily accurate to assume that a level measured as "comfortable" using one method will be the same as a level also labeled "comfortable" obtained using a different method. This can readily be demonstrated by comparing the data in Table 2 with data reported by Allen, Hall, and Jeng (1990) for a category rating loudness measurement procedure that differs in several ways from the Contour Test. Despite their differences, both procedures include categories labeled "very soft," "comfortable/OK," and "loud." Table 5 compares the mean sound pressure levels corresponding to each of these categories for a 1000 Hz stimulus for normal-hearing listeners. Although the level corresponding to "loud" is quite similar for the two procedures, those corresponding to "very soft" and "comfortable" are considerably different. Because all of the subjects had normal hearing, it is likely that these sizable differences in measured levels corresponding to loudness categories are due to procedural effects rather than to sampling error. Clearly, these differences would have to be accounted for in the application of the loudness data from either test for hearing aid prescription purposes.

	Very Soft	Comfortable	Loud
Allen et al. (1990)	39	83	97
This study	21	89	95

TABLE 5. Sound pressure levels corresponding to three loudness category descriptors measured by two procedures. Data are for normal-hearing listeners at 1000 Hz.

Accounting for Loudness Summation Across Bandwidth [TOP](#)

To generate the data required for hearing aid frequency response shaping, loudness category levels must be measured using stimuli with restricted bandwidth. However, because of the effects of loudness summation across bandwidth, caution must be exercised when applying these results to the prescription of required gain for different input levels. Figure 5 shows quite clearly that the loudness growth function is different for speech (a sound of complex and varying bandwidth) than for warble tones. Although these data are for normal-hearing listeners, it is very likely that differences also would be observed for hearing-impaired listeners. Thus, it cannot be assumed that a particular sound pressure level that is comfortable when presented as a narrowband stimulus will also be comfortable when presented as a broader band stimulus. The hearing aid prescription strategy must account for this issue in some way. Unfortunately, the magnitude of loudness summation across bandwidth for listeners with sensorineural hearing loss is uncertain. Thus, appropriate strategies can only be postulated at this time. The VIOLA procedure addresses this issue by attempting to replicate, with amplification, the relationship that is observed in normal-hearing listeners between typical speech levels and the loudness of warble tones. This assumes that hearing-impaired listeners experience about the same loudness summation across bandwidth at different stimulus levels as do normal-hearing listeners. Some investigations support this assumption (Denley-McFtridge, Reference Note 5; Khalil, Skinner, Miller & Andrews, Reference Note 7), whereas others suggest that hearing-impaired listeners experience less (Bonding, 1979) or more (Bentler & Pavlovic, 1989) loudness summation across bandwidth than normal-hearing individuals. The problem is further complicated by the fact that loudness summation across bandwidth varies with level, thus, normal-hearing and hearing-impaired listeners might be similar at some levels but different at other levels. Finally, it is possible that loudness summation patterns differ across hearing-impaired persons just as loudness growth functions do. If this is the case, it might be necessary to develop a method of clinically measuring loudness summation across bandwidth as a function of level and frequency to refine hearing aid prescription procedures.

Conclusion [TOP](#)

The Contour Test was developed specifically for use as a clinical method to quantify loudness perceptions of hearing-impaired listeners for use in hearing aid prescription and selection. This paper reports norms for the test and provides equations that can be used to generate a template to compare responses from patients to those of typical normal-hearing listeners. The test procedure, which consumes about five minutes per test frequency, appears to be almost as reliable as a standard threshold test (e.g., Byrne & Dillon, 1981) and about equal in reliability to the few comparable tests of loudness perception that have been reported.

The recently increased prevalence of FDRC hearing aids holds forth the promise of substantially improved benefit when properly applied, while simultaneously presenting hearing aid specialists with new challenges in matching the technology to the needs of patients. As the profession moves in the direction of developing clinical procedures to respond to this challenge, there is a need to agree on, and adopt, a standard procedure for quantifying loudness perceptions. The procedure must satisfy the scientific requirements of producing valid and reliable data as well as the clinical requirements of appropriate cost/benefit ratio and good patient acceptance. The strengths of the Contour Test are in its flexible administration (manual or computer-controlled), good patient acceptance, adequate reliability, and availability of normative data. Other matters such as the validity of the data yielded by the psychophysical procedure and the cost/benefit ratio for the practitioner can only be determined on the basis of clinical applications and experience with hearing aid fittings that have used the procedure.

Acknowledgments: [TOP](#)

Supported by funding from the Department of Veterans Affairs Rehabilitation Research and Development Service. Christy Goff, Sara Martin and Lori McCloud assisted with data collection.

References [TOP](#)

- Allen, J. B., Hall, J. L., & Jeng, P. S. (1990). Loudness growth in ½-octave bands (LGOB): A procedure for the assessment of loudness. *Journal of the Acoustical Society of America*, *88*, 745-753.
[\[Context Link\]](#)
- American National Standards Institute (1978). S3.21:"*Methods for Manual Pure Tone Threshold Audiometry.*" New York: American National Standards Institute.
[\[Context Link\]](#)
- American National Standards Institute (1989). S3.6:"*Specification for Audiometers.*" New York: American National Standards Institute.
[\[Context Link\]](#)
- Beattie, R. C., Svihovic, D. A., Carmen, R. E., & Kunkel, H. A. (1980). Loudness discomfort level for speech: Comparison of two instructional sets for saturation sound pressure level selection. *Ear and Hearing*, *1*, 197-205.
[\[Context Link\]](#)
- Bentler, R. A., & Pavlovic, C. V. (1989). Comparison of discomfort levels obtained with pure tones and multitone complexes. *Journal of the Acoustical Society of America*, *86*, 126-132.
[\[Context Link\]](#)
- Bonding, P. (1979). Critical bandwidth in loudness summation in sensori-neural hearing loss. *British Journal of Audiology*, *13*, 23-30.
[\[Medline Link\]](#) [\[CrossRef\]](#) [\[Context Link\]](#)
- Byrne, D. (1986). Effects of bandwidth and stimulus type on most comfortable loudness levels of hearing-impaired listeners. *Journal of the Acoustical Society of America*, *80*, 484-493.
[\[Context Link\]](#)
- Byrne, D. (1996). Hearing aid selection for the 1990s: Where to? *Journal of the American Academy of Audiology*, *7*, 377-395.
[\[Context Link\]](#)
- Byrne, D., & Dillon, H. (1981). Comparative reliability of warble tone thresholds under earphones and in sound field. *Australian Journal of Audiology*, *3*, 12-14.
[\[Context Link\]](#)
- Carhart, R., & Jerger, J. F. (1959). Preferred method for clinical determination of pure tone thresholds. *Journal of Speech and Hearing Disorders*, *24*, 330-345.
[\[Context Link\]](#)
- Cox, R. M. (1989). Comfortable loudness level: Stimulus effects, long-term reliability and predictability. *Journal of Speech and Hearing Research*, *32*, 816-828.
[\[Context Link\]](#)
- Cox, R. M. (1995). Using loudness data for hearing aid selection: The IHAFF approach. *Hearing Journal*, *47*(2), 10, 39-42.
[\[Context Link\]](#)
- Cox, R. M., Alexander, G. C., Gilmore, C. A., & Pusakulich, K. M. (1989). The Connected Speech Test version 3: Audiovisual administration. *Ear and Hearing*, *10*, 29-32.
[\[Context Link\]](#)
- Elberling, C., & Neilsen, C. (1993). The dynamics of speech and the auditory dynamic range in sensorineural hearing impairment. In J. Beilin & G. R. Jensen (Eds). *Recent developments in hearing instrument technology* (pp. 99-134). Denmark: Danavox Jubilee Foundation.
[\[Context Link\]](#)
- Filion, P., & Margolis, R. (1992). Comparison of clinical and real-life judgments of loudness discomfort. *Journal of the American Academy of Audiology*, *3*, 193-199.
[\[Context Link\]](#)
- Florentine, M., Buus, S., & Poulsen, T. (1996). Temporal integration of loudness as a function of level. *Journal of the Acoustical Society of America*, *99*, 1633-1644.

[\[Context Link\]](#)

Geller, D., & Margolis, R. (1984). Magnitude estimation of loudness I: Application to hearing aid selection. *Journal of Speech and Hearing Research*, 27, 20-27.

[\[Context Link\]](#)

Grey, H., & Dyrland, O. (1996). Putting loudness scaling to work: Moving from theory to practice. *The Hearing Journal*, 49(3), 49-54.

[\[Context Link\]](#)

Hawkins, D. B. (1980). The effect of signal type on the loudness discomfort level. *Ear and Hearing*, 1, 38-41.

[\[Context Link\]](#)

Hawkins, D. B., Walden, B. E., Montgomery, A., & Prosek, R. A. (1987). Description and validation of an LDL procedure designed to select SSPL90. *Ear and Hearing*, 8, 162-169.

[\[Context Link\]](#)

Hellbruck, J., Thomamuller, D., & Zeitler, A. (1995). Hearing field audiometry with groups of normal hearing persons in natural sound fields. *Audiologische Akustik*, 34, 60-81.

[\[Context Link\]](#)

Jesteadt, W., Luce, R. D., & Green, D. M. (1977). Sequential effects in judgment of loudness. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 92-104.

[\[Context Link\]](#)

Kiessling, J., Schubert, M., & Archut, A. (1996). Adaptive fitting of hearing instruments by category loudness scaling(Scal-Adapt). *Scandinavian Audiology*, 25, 153-160.

[\[Context Link\]](#)

Kiessling, J., Steffens, T., & Wagner, I. (1993). On the clinical applicability of loudness scaling. *Audiologische Akustik*, 32, 100-115.

[\[Context Link\]](#)

Mellers, B. A. (1983). Evidence against "absolute" scaling. *Perception and Psychophysics*, 33, 523-526.

[\[Context Link\]](#)

Orchik, D. J., & Mosher, N. L. (1975). Narrow band noise audiometry: The effect of filter slope. *Journal of the American Audiology Society*, 1, 50-53.

[\[Context Link\]](#)

Pascoe, D. P. (1986). Hearing aid selection procedure used at Central Institute for the Deaf in St. Louis. *Audiological Acoustics*, 25, 90-106.

[\[Context Link\]](#)

Ricketts, T. A. (1996). Fitting hearing aids to individual loudness-perception measures. *Ear and Hearing*, 17, 124-132.

[\[Context Link\]](#)

Ricketts, T. A., & Bentler, R. A. (1996). The effect of test signal type and bandwidth on the categorical scaling of loudness. *Journal of the Acoustical Society of America*, 99, 2281-2287.

[\[Context Link\]](#)

Ricketts, T. A., & Van Vliet, D. (1996). Updating our fitting strategies given new technology. *American Journal of Audiology*, 5, 29-35.

[\[Context Link\]](#)

Robinson, K., & Gatehouse, S. (1996). Test-retest reliability of loudness scaling. *Ear and Hearing*, 17, 120-123.

[\[Context Link\]](#)

Sammeth, C. A., Birman, M., & Hecox, K. E. (1989). Variability of most comfortable and uncomfortable loudness levels to speech stimuli in the hearing impaired. *Ear and Hearing*, 10, 94-100.

[\[Context Link\]](#)

Ventry, I. M., & Johnson, J. I. (1978). Evaluation of a clinical method for measuring comfortable loudness for speech. *Journal of Speech and Hearing Disorders*, 43, 149-159.

[\[Context Link\]](#)

Ward, L. M., & Lockhead, G. R. (1970). Sequential effects and memory in category judgments. *Journal of Experimental Psychology*, 84, 27-34.

[\[CrossRef\]](#) [\[Context Link\]](#)

Woods, R. W., Ventry, I. M., & Gatling, L. W. (1973). Effect of ascending and descending measurement methods on comfortable loudness levels for pure tones. *Journal of the Acoustical Society of America*, 54, 205-206.

[\[Context Link\]](#)

Zwicker, E., Flottorp, G., & Stevens, S. S. (1957). Critical band width in loudness summation. *Journal of the Acoustical Society of America*, 29, 548-557.

[\[Context Link\]](#)

Reference Notes [TOP](#)

1 Cornelisse, L. E., & Jamieson, D. G. (1994). Advanced measures of residual auditory capacity using an audiometer under computer control. Presented at the National Convention of the American Academy of Audiology, Richmond, VA.

[\[Context Link\]](#)

2 Cox, R. M., Goff, C. M., Martin, S. E., & McCloud, L. L. (1994). *The Contour Test: Normative data*. Presented at the National Convention of the American Academy of Audiology, Richmond, VA.

[\[Context Link\]](#)

3 Cox, R. M., Taylor, I. M., Gray, G. A., & Brainerd, L. E. (1994). *The Contour Test: Applications to hearing aid selection and fitting*. Presented at the National Convention of the American Academy of Audiology, Richmond, VA.

[\[Context Link\]](#)

4 Cox, R. M., & Alexander, G. C. (1995). *The relationship between loudness contours and preferred listening levels*. Presented at the National Convention of the American Academy of Audiology, Dallas, TX.

[\[Context Link\]](#)

5 Denley-McFatrige, L. K. (1990). *Loudness summation across bandwidth: Effect of hearing loss and psychophysical procedure*. Unpublished Master's Thesis, Memphis State University, Memphis, TN.

[\[Context Link\]](#)

6 Jenstad, L. M., Cornelisse, L. E., & Seewald, R. C. (1996). *How are you measuring loudness growth? It does matter*. Presented at the National Convention of the American Academy of Audiology, Salt Lake City, UT.

[\[Context Link\]](#)

7 Khalil, E. H., Skinner, M. W., Miller, J. D., & Andrews, S. (1989). *Monaural and binaural loudness summation of speech*. Unpublished manuscript, Central Institute for the Deaf, St. Louis, MO.

[\[Context Link\]](#)

8 Knier, E. C., & Bentler, R. A. (1994). *Development of a clinical method for determining loudness perception*. Presented at the National Convention of the American Academy of Audiology, Richmond, VA.

[\[Context Link\]](#)

9 Nielsen, J. R. (1995). *Loudness scaling: Influence from some experimental parameters*. Technical University of Denmark. Acoustics Laboratory Report no. 65.

[\[Context Link\]](#)

*To give the patient a context to decide whether a sound is uncomfortably loud, it can be useful to review the kinds of sounds (especially music) that they listen to. Most patients experience some occasions, such as church music, concerts, theater, dances, etc., when they enjoy fairly loud sounds. [\[Context Link\]](#)

†The Contour test can be administered and scored under computer control using an audiometer with a serial interface. Software to administer the test on many audiometers can be obtained by contacting the first author or through an inquiry to the Hearing Aid Research Laboratory page on the Internet (<http://www.ausp.memphis.edu/harl>). The test can also be given using manual administration. Score sheets for manual testing, instructions, a category listing similar to [Figure 1](#), and a template suitable for making clinical comparisons of patient data with norms also can be downloaded from the Internet address given above. [\[Context Link\]](#)

© Williams & Wilkins 1997. All Rights Reserved.

Copyright © 2008, Lippincott Williams & Wilkins. All rights reserved.

Published by Lippincott Williams & Wilkins.

[Copyright/Disclaimer Notice](#) • [Privacy Policy](#)

 [Subscribe to our RSS feed](#)

txrdc-pt01.tx.ovid.com
Release 4.7.0