Veterans Administration

Journal of Rehabilitation Research and Development Vol. 24 No. 4
Pages 283–292

Towards a general measure of distortion*

HARRY LEVITT, ED CUDAHY, WEI-HUA HWANG, ELIZABETH KENNEDY, and CYNTHIA LINK**

Center for Research in Speech and Hearing Sciences, Graduate School and University Center, City University of New York, New York 10036 and Lexington Center, Inc., Jackson Heights, New York 11370

Abstract—An approach to the measurement of distortion in hearing aids is presented. Distortion is first defined in general terms. The technique is then applied to the specific problem of predicting the detectability of non-linear distortion. Distortion indices have been obtained for two common, but very different forms of distortion: peak clipping (common to most conventional hearing aids); and, quantization (a form of distortion common to digital systems including digital hearing aids). Distortion indices were derived for the two forms of distortion in steady-state vowels. The results showed good agreement between the distortion indices at the just-detectable level of distortion.

INTRODUCTION

Current techniques for measuring distortion in hearing aids are geared primarily to the measurement of harmonic distortion. Whereas percent harmonic distortion can be measured with great precision for tonal signals, these physical measurements provide only a rough indication of the perceptual interference produced by harmonic distortion. Further, measurements of harmonic distortion with tonal signals provide very little information on how this form of distortion affects more complex signals, such as speech. A particularly troublesome problem is that of intermodulation distortion, in which harmonic distortion components interact with each other and with related frequency components in the signal. Even if a satisfactory technique were to be developed for measuring and specifying intermodulation distortion, such a solution would address only one problem among many.

There is a pressing need to develop a general measure of distortion. It is important that this general measure of distortion be applicable not only to those forms of distortion produced by conventional hearing aids, but also to new forms of distortion that are likely to be encountered in the digital hearing aids of the future. Existing methods of measuring distortion are inadequate even for conventional hearing aids, and there are no established methods for specifying or measuring digitally generated distortions in hearing aids.

The purpose of this paper is to develop a general approach to the measurement and specification of distortion in audio systems, and in hearing aids in particular. The first step in this approach is to define distortion in mathematically tractable terms. The second stage involves the development of a practical procedure for identifying and measuring distortion in an amplified audio signal. In the third stage, criteria are established for predicting important perceptual aspects of the measured distortion (e.g., audibility or acceptability of the perceived distortion).

*This work was supported by Rehabilitation Engineering Center Grant No. G00832511, from the National Institute of Disability and Rehabilitation Research (NIDRR) to the Lexington Center, Inc.

**Harry Levitt is with the Center for Research in Speech and Hearing Sciences. All other authors are with the Lexington Center, Inc.

Journal of Rehabilitation Research and Development Vol. 24 No. 4 Fall 1987

DEFINITION OF DISTORTION

A fundamental problem in developing a general measure of distortion is that of defining distortion in mathematically tractable terms. This problem reduces essentially to that of separating changes produced by the amplifying system that are desirable (e.g., increased intensity, frequency shaping to improve intelligibility or quality) from changes that are undesirable. In general, distortion is defined as those changes produced by an amplifying system that are undesirable.

It is very useful in this context to draw a distinction between distortions produced by linear and non-linear operations on the audio signal. Non-linear operations (e.g., clipping) almost invariably have an undesirable effect on the quality of the audio signal. Since it is possible to separate linear from non-linear operations in precise mathematical terms, this distinction provides a means for specifying unambiguously a very large class of distortions.

A linear system is defined as follows. Consider a system in which an input signal $a_{in}(t)$ produces an output response $a_{out}(t)$. For a second input, $b_{in}(t)$, the response will be $b_{out}(t)$. If the response of the system to the sum of the two input signals $a_{in}(t) + b_{in}(t)$ is equal to $a_{out}(t) + b_{out}(t)$, then the system is said to be linear. In the case of a non-linear system, the summed output is $ab_{out}(t)$, where $ab_{out}(t)$ is not equal to $a_{out}(t) + b_{out}(t)$.

A very useful property of linear systems is that the output can be derived explicitly from a knowledge of the input and the system's response to a single impulse; i.e.,

$$a_{out}(t) = a_{in}(t) * h(t) \qquad [1]$$

where h(t) is the impulse response of the system, and $*$ indicates convolution.

Since convolution is difficult to compute (and for non-mathematicians, difficult to comprehend) it is common practice to take the Fourier transform of both sides of Equation 1. This yields the more familiar equation:

$$A_{out}(f) = A_{in}(f) \,.H(f) \qquad [2]$$

where $A_{out}(f)$ is the frequency spectrum of $a_{out}(t)$,
$A_{in}(f)$ is the frequency spectrum of $a_{in}(t)$, and
H(f) is the frequency response of the linear system.

Another useful property of a linear system is that the output $A_{out}(f)$ will contain the same frequency components as the input $A_{in}(f)$. Thus, if the input to a linear system consists of a single frequency component at fo, the output will also consist of a single frequency component at fo. The only difference between input and output will be a change in the amplitude and phase of this frequency component. In contrast, the output of a non-linear system will contain frequency components that are not contained in the input to the system.

In general, the output of a non-linear system for a single frequency input $A_{in}(f_i)$ is given by

$$A_{out}(f) = A_{out}(f_i) + \sum_{j \neq i} A_{out}(f_j) \qquad [3]$$

The above property serves as the basis for defining non-linear distortion. For our purposes, the distortion spectrum for a non-linear system is defined as:

$$D(f) = AB_{out}(f) / [A_{out}(f) + B_{out}(f)] \qquad [4]$$

where D(f) is the distortion spectrum,
$AB_{out}(f)$ is the spectrum of the output of the non-linear system for an input $[A_{in}(f) + B_{in}(f)]$, and
$[A_{out}(f) + B_{out}(f)]$ is the spectrum of the output of the idealized linear system for the same input.

The specification of the idealized linear system can present a problem. The easiest situation to deal with is one in which the properties of the idealized linear system are known in advance. In this case, the input and output signals, $A_{in}(f)$, $B_{in}(f)$ and $AB_{out}(f)$, can be measured directly, and the idealized output $A_{out}(f) + B_{out}(f)$ can be derived without difficulty. The distortion spectrum D(f) can then be determined directly from Equation 4.

The difficult situation is one in which the idealized linear system is not known in advance. It is possible, however, to derive an idealized linear system that provides a good approximation to the nonlinear system being evaluated in the following way. A single frequency input, $A_{in}(fi)$, is used and the output at frequency f_i only is measured. The resulting output $A_{out}(fi)$ is assumed to be that produced by the idealized linear system. In this way, the idealized linear system can be derived, i.e.,

$$Ho(fi) = A_{out}(fi) / A_{in}(fi) \qquad [5]$$

where Ho(f) is the frequency response of the idealized linear system.

Given this information it is a relatively simple matter to determine the distortion term D(f) in Equation 4. Note that the value of $\sum_{j \neq} A_{out}(f_j)$ in Equation 3 will depend on the magnitude of the input signal and that D(f) will vary as a function of signal level.

It is a much more difficult problem to define distortion in a linear system in mathematically tractable terms. For example, the frequency-gain characteristic of a properly prescribed hearing aid may represent a desirable change in the audio signal and consequently the output will not be distorted according to the general definition of distortion presented earlier. On the other hand, the frequency-gain characteristic of an improperly prescribed hearing aid is likely to produce undesirable changes in the audio signal. Under such conditions, the hearing aid user will typically report that the amplified signal sounds distorted. Similarly, irregularities (e.g., ripple) in the frequency-gain characteristic may also cause the amplified signal to sound distorted. In both of these cases the operation of the hearing aid may be linear in a mathematical sense, but the result is nevertheless a distorted signal.

A related problem is the distortion produced by room reverberation. Sound produced in an anechoic room will sound unnatural. The addition of multiple acoustic reflections, as occurs in a typical room, improves the quality of the sound. Too many reflections, or reflections with very long delays, can distort the audio signal beyond recognition, although all of the operations involved may be strictly linear.

If the idealized linear system is known, then distortion can be defined quite simply as:

$$D(f) = A1_{out}(f) / Ao_{out}(f) \qquad [6]$$

where $A1_{out}(f)$ is the signal spectrum at the output of the system being evaluated, and
$Ao_{out}(f)$ is the spectrum of the signal at the output of the idealized linear system,

Using Equation 2, the above can be rewritten as:

$$D(f) = A_{in}(f) \,.H1(f) / A_{in}(f) \,.Ho(f) \qquad [7]$$

where $A_{in}(f)$ is the spectrum of the input signal, which is common to both systems,
H1(f) is the frequency response of the linear system being evaluated, and
Ho(f) is the frequency response of the idealized linear system.

Note that $A_{in}(f)$, $A1_{out}(f)$, $Ao_{out}(f)$, H1(f), and Ho(f) are complex quantities and are typically specified in terms of both amplitude and phase. It should also be noted that Equation 7 is applicable only to linear systems, whereas Equation 6 is quite general and can be used with both linear and non-linear systems. Equation 6 may be viewed as a more general form of Equation 4 and is used here as a general definition of distortion.

A difficult problem is that of defining linear distortion when the idealized linear system is not known. The approach used here is to determine first those changes produced by the amplifying system which are just detectable and then to determine the extent to which these changes are acceptable or unacceptable to the user. The purpose of this paper is to introduce the mathematical framework for predicting just detectable differences produced by an amplifying system, whether linear or nonlinear. This mathematical framework will be extended in subsequent work to cover other issues, such as acceptability of the distortion.

A GENERAL MEASURE OF DISTORTION

Distortion in an audio system can range from a level that is just audible to a level that renders the signals processed by the system to be totally unrecognizable. This paper is concerned only with just detectable levels of distortion. This level of distortion is easy to define and can be measured with relative precision. It also represents a lower bound in determining acceptable levels of distortion; i.e., distortion that is inaudible is unlikely to be of any practical concern.

Of particular interest is the development of a predictive formula for the just audible level of distortion. The proposed measure of distortion is based on Equation 6, for deriving D(f), suitably modified to take into account the characteristics of the human auditory system. Specifically, it is assumed that the perception of auditory distortion involves three distinct operations:

1) the signals reaching the ear are analyzed by a bank of critical-band filters;

2) in order to take spread-of-masking effects into account, the critical-band spectrum obtained in #1 above is replaced by an *effective critical-band spec-*

trum, in which large variations between critical bands have been smoothed out; and

3) the distortion energy within each critical-band filter is averaged over a period of time corresponding to the integration time of the ear. Distortion energy is defined by analogy with Equation 6 using the effective critical-band spectra in place of $Ao_{out}(f)$ and $A1_{out}(f)$.

It is recognized that critical bandwidths, integration times and spread-of-masking effects will vary between individuals and that there are likely to be large differences in these variables between normal-hearing and hearing-impaired listeners. At the same time, there is a need to develop a distortion measure that can be derived from physical measurements of the system being evaluated and not be dependent on psychophysical measurements of the person using the system. As a consequence, two distortion measures are proposed. The first is derived solely from physical measurements with typical normative values being used for critical bandwidths, integration times, and spread-of-masking effects. This measure, the Distortion Index, is defined as:

where DI(i) is the Distortion Index for critical band i
$ECB_1(i)$ is the effective critical band spectrum at the output of the system being evaluated, and
$ECB_0(i)$ is the effective critical band spectrum at the output of the idealized linear system.

Note that $ECB_1(i)$ and $ECB_0(i)$ are expressed in decibels; thus the difference between the two spectra is used rather than their ratio.

The second measure, the Distortion Index for Individuals, DII(i, j), is derived in the same way as DI(i) except that the critical bandwidths, integration times, and spread-of-masking effects are measured separately for each individual. The suffix j identifies the individual for whom the distortion measure has been derived. As before, the suffix i identifies the critical band.

Of the two measures, DI(i) is of primary interest since it represents a measure that can be implemented in a practical instrument. DII(i, j) is of value as a research tool for investigating the factors affecting the perception of distortion.

In order to derive DI(i) it is necessary to assume normative values for critical bandwidths, integration times, and spread-of-masking effects. Scharf (11) provides a comprehensive review of published research on the critical band, and the average critical bandwidths derived by Scharf are used here.

Research on the temporal characteristics of the ear indicates that two integration times need to be considered. The first relates to the time interval within which a short-term spectral analysis of the incoming sound is performed. Typical values from the published literature indicate a time window on the order of 10 to 15 ms (8). The second integration time relates to the summation of acoustic energy by the auditory system in the detection of short-duration stimuli. Typical values of this integration time range from 200 to 400 ms (13,15). The integration times used in this study were 12.8 and 384 ms respectively; the latter integration time corresponds to 30 time windows of 12.8 ms each.

Spread-of-masking effects have been studied by Kryter (5), Bilger and Hirsh (2), and others. The most important of these effects is an upward spread-of-masking. Kryter provides a simplified set of rules for approximating spread-of-masking effects. For signals at a comfortable listening level, upward spread-of-masking is on the order of 24 dB/octave. Since the critical bandwidth is about 1/6 octave wide (above 600 Hz), the effective critical-band spectrum of a sound will show a maximum rate of fall of 4 dB between bands for band level decreasing with frequency increasing. For example, if the sound level in one critical band is 75 dB SPL and the sound level in the next highest critical band is 69 dB SPL, then the effective level in the second critical band will be 75 − 4 = 71 dB SPL.

The derivation of DI(i) in a practical system is illustrated in **Figures 1** through **5**. The system was configured around an IBM XT personal computer. The audio signal was sampled at a rate of 20,000 samples/second and converted to digital form with 12-bit resolution using a Data Translation 2801A analog-to-digital converter board. A fast Fourier transform was used to obtain the short-term spectrum of both the distorted and undistorted waveforms. A time window of 12.8 ms (= 256 samples at 20 kHz) was used for the spectrum analysis. A typical pair of spectra is shown in **Figure 1**. The undistorted spectrum is represented by open squares, the distorted spectrum by open diamonds. The data are for a single time window during the vowel /i/, as in *seen*. Peak clipping of the time waveform was

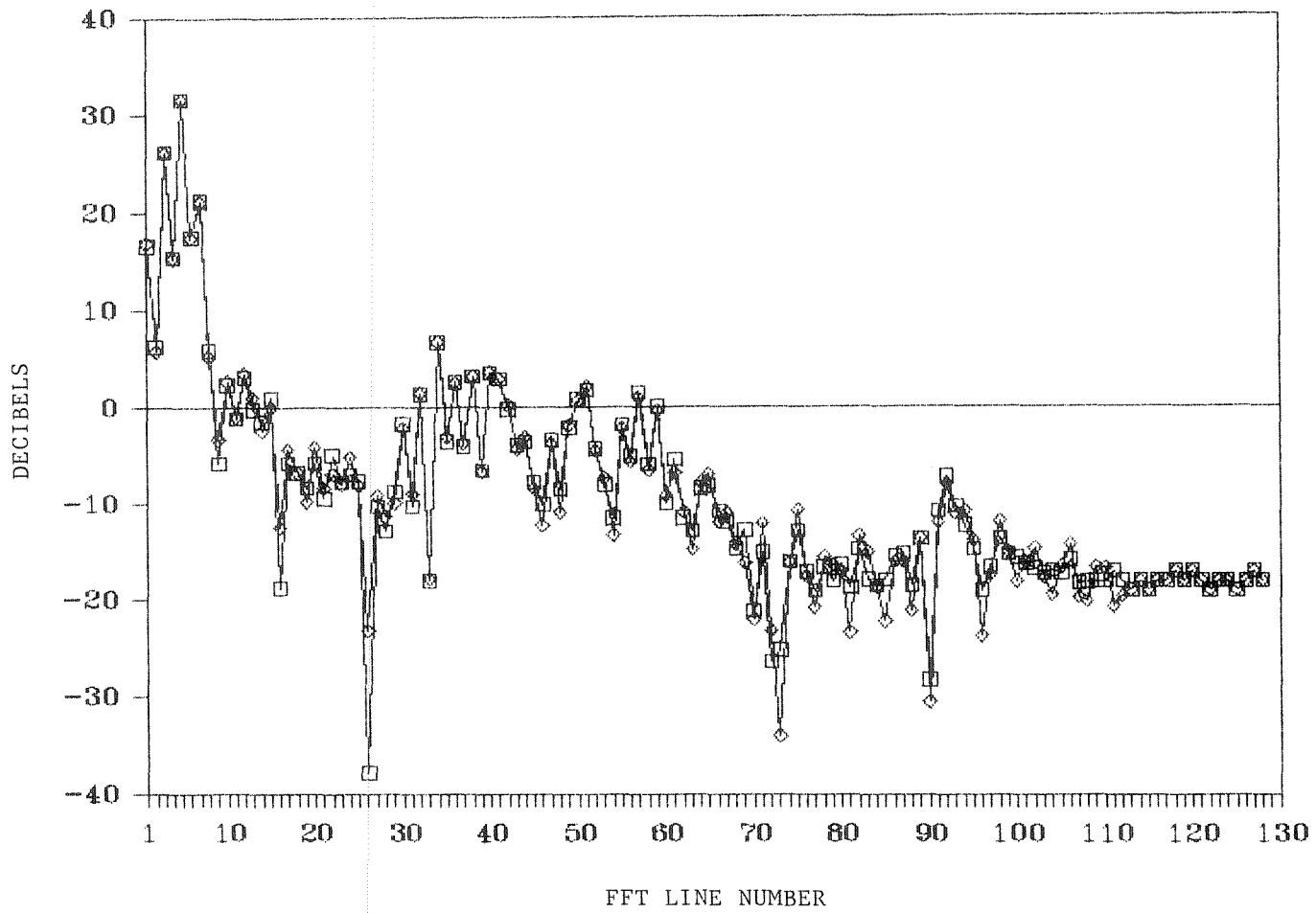

Figure 1.
Short-Term Power Spectra. Short-term power spectra are shown for the vowel /i/ for a time window of 12.8 ms. The undistorted spectrum is shown by open squares (□), the spectrum of the distorted signal (2-percent peak clipping) is shown by open diamonds (◇). A line spectrum is shown since a discrete Fourier transform was used. Each spectrum line corresponds to 78.125 Hz.

the form of distortion used in this example. The waveform was clipped 2 percent of the time, which was the just detectable level of peak clipping for this vowel.

The short-term vowel spectrum shown in **Figure 1** has major peaks at frequencies of 390, 2656, and 3984 Hz, respectively. These peaks correspond to the first three formants of the vowel. The vowel was produced by a young adult female and these formant values are typical of those reported by Peterson and Barney (9) for female talkers.

The spectra shown in **Figure 1** were converted to critical-band spectra using the critical bandwidths shown in **Table 1**. The critical-band spectra were then smoothed allowing for a 24dB/octave upward spread of masking. **Figure 2** shows the critical-band spectra, prior to and after smoothing, for the undistorted spectrum of **Figure 1**. Note that the unsmoothed critical-band spectrum shows a marked drop in level in going from Band #3 to Band #5. The effective (smoothed) critical-band spectrum shows a steady decline in level of 4 dB/octave from Band #3 through Band #11. The smoothing operation is essentially the result of an upward spread of masking that occurs in the valley between the first two formants.

The distortion index for a single time window was derived by computing the difference (in dB) between the effective critical-band spectrum of the distorted signal and the effective critical-band spectrum of the undistorted signal. The distortion index for a single time window is referred to as $DI_k(i)$ where the suffix $_k$ identifies the time window. **Figure 3** shows the value of $DI_k(i)$ as a function of i for time

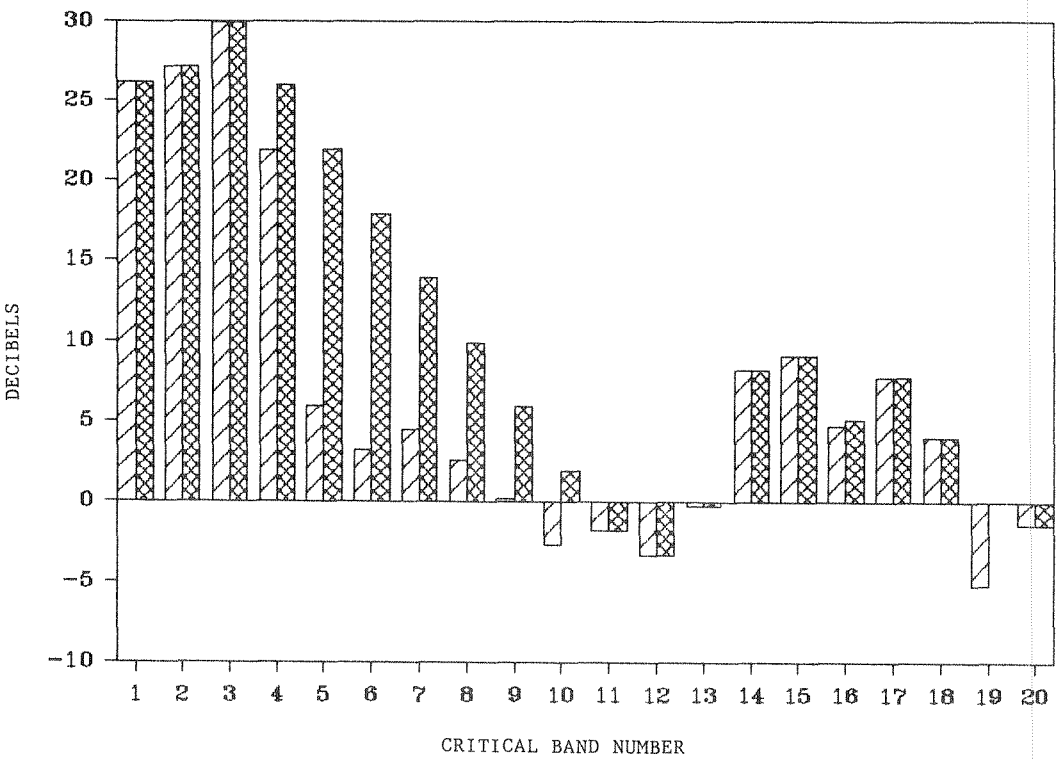

Figure 2.
Critical-Band Spectra. Critical band spectra are shown for the undistorted spectrum of **Figure 1.** The single-hatched bars show the original (unsmoothed) critical-band spectrum. The double-hatched bars show the effective (smoothed) critical-band spectrum.

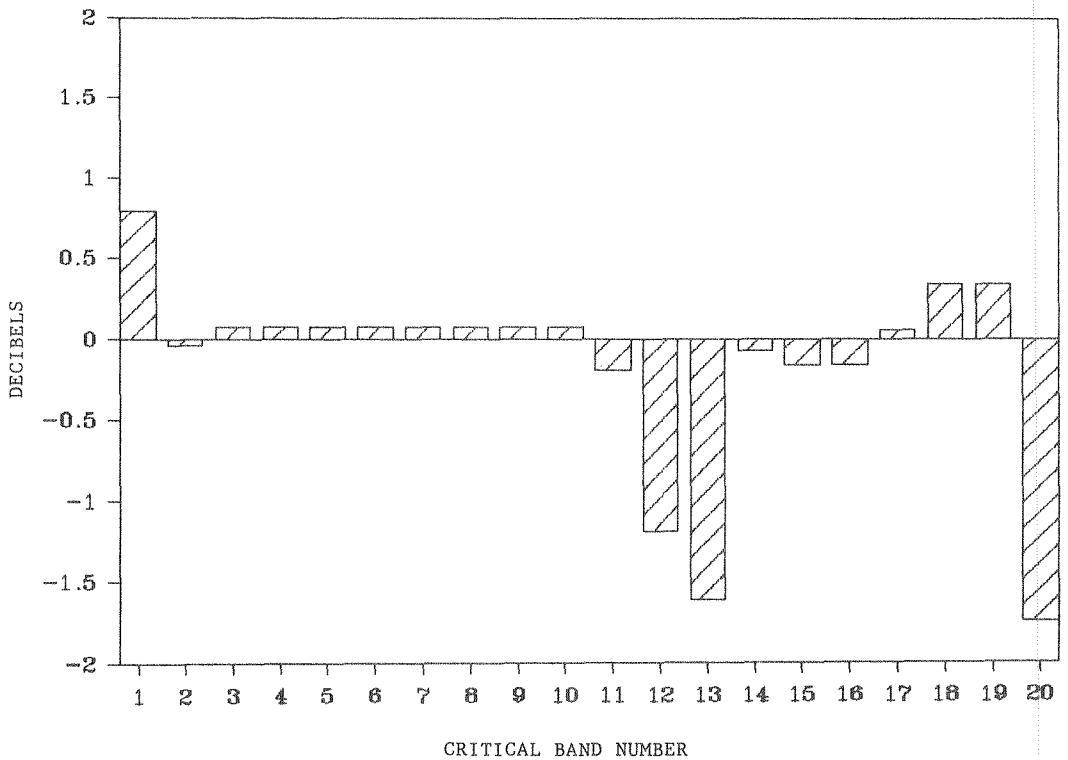

Figure 3.
Critical-Band Distortion Spectrum for a Single Time Window. The histograms show the difference between the effective critical-band spectrum for the distorted signal and the effective critical-band spectrum for the undistorted signal. The data are for Time Window #15 of the vowel /i/. The distorted signal has been peak clipped for 2-percent of the time waveform.

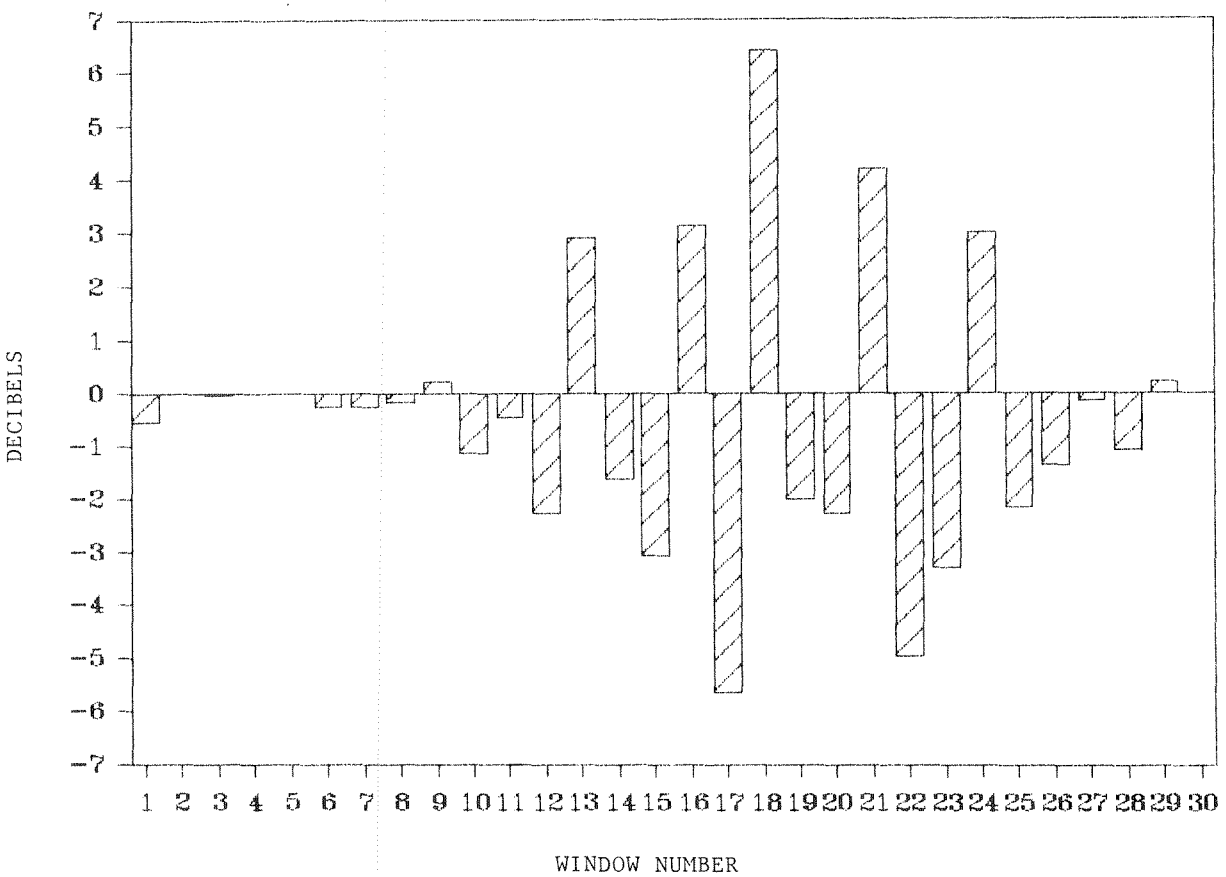

Figure 4.
Variation in Distortion Level with Time. Distortion level for Band #13 is shown for successive time windows of the vowel /i/. Distortion is 2-percent peak clipping.

Table 1.
Critical Bandwidths (From Scharf, 1970 [11]).

Band #	Critical Band cf	Lower cutoff	Upper cutoff
1	150	100	200
2	250	200	300
3	350	300	400
4	450	400	510
5	570	510	630
6	700	630	770
7	840	770	920
8	1000	920	1080
9	1170	1080	1270
10	1370	1270	1480
11	1600	1480	1720
12	1850	1720	2000
13	2150	2000	2320
14	2500	2320	2700
15	2900	2700	3150
16	3400	3150	3700
17	4000	3700	4400
18	4800	4400	5300
19	5800	5300	6400
20	7000	6400	7700

window #15. Note that **Figures 1, 2, and 3** all relate to the same time window (k = 15).

The largest value of $DI_k(i)$, in absolute magnitude, occurs in Band #20 for the time window shown. When averaged over the 30 time windows considered in the analysis, Band #13 showed the largest average difference. **Figure 4** shows the variation in DI_k (13) as a function if k; i.e., the diagram shows the variation in distortion level for Band #13 for successive time windows. It is evident from the figure that DI_k (13) varies substantially with time.

Figure 5 shows the mean values of DI(i) averaged over k (k = 1,2, . . . 30). In deriving this average, the distortion energy was first derived for each time window and critical band. The total energy over the 30 time windows was obtained for each critical band and then the average power in decibels was computed for each critical band. The largest value of

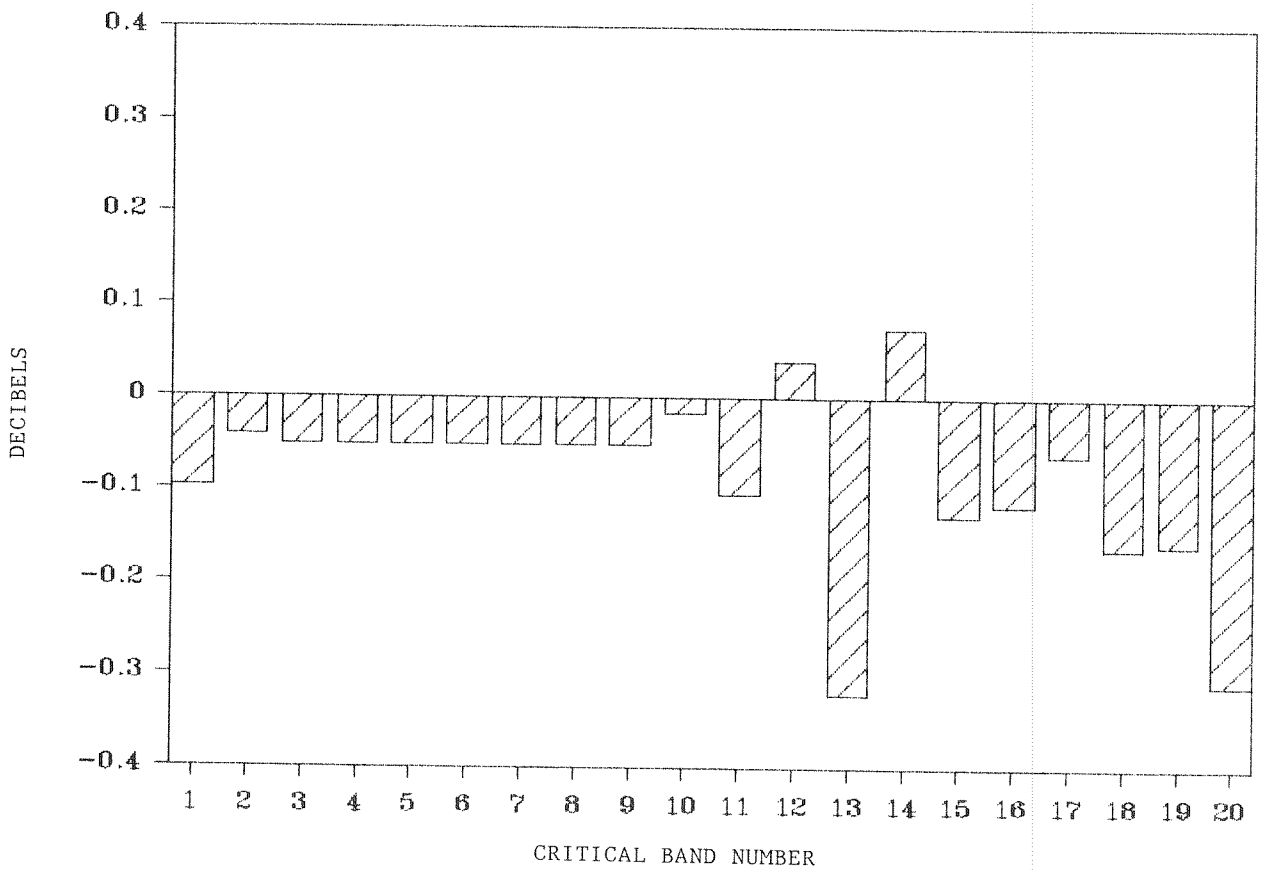

Figure 5.
Distortion Index DI(i). The Distortion Index DI(i) is shown for the vowel /i/ subject to 2-percent peak clipping. The distortion index has been averaged over 30 time windows. $DI_{max}(i)$ has a value of 0.3 dB (absolute value) for Band #13.

DI(i) is referred to as $DI_{max}(i)$. As can be seen from the figure, $DI_{max}(i)$ is 0.3 dB and occurs in Band #13. Note that DI(i) for Band #20 is only slightly smaller in magnitude than that for Band #13. In a revised version of the model (now under consideration), the rule for detection is modified when the distortion energies in two or more critical bands are identical.

INTERPRETATION OF THE DISTORTION MEASURE

DI(i) represents a series of indices, one for each critical band. It is assumed that distortion is just audible when DI(i) exceeds some critical value in any one critical band. **Table 2** shows estimates of this critical value as obtained for three vowel sounds at the just audible level of distortion for two types of distortion, quantization, and clipping. Two sets of data are shown: one in which an upward spread of masking of roughly 4 dB/octave is assumed (the left half of **Table 2**), and a second data set in which spread of masking is not taken into account (the right half of **Table 2**). Each set of data shows the value of $DI_{max}(i)$ at which distortion is just audible, the critical band (i) in which DI(i) is a maximum (averaged over 30 time windows), and the center frequency of the critical band.

The critical value of $DI_{max}(i)$ at which distortion is just audible is on the order of half a decibel, on the average. Note that the just detectable value of $DI_{max}(i)$ varies over a wider range for quantization distortion than that for peak clipping. It should be remembered, however, that quantization distortion was varied in relatively large steps (1 bit per step)

Table 2.
Values of $DI_{max}(i)$.

Sound Type and Form of Distortion	Smoothed Critical-Band Spectra			Unsmoothed Critical-Band Spectra		
	$DI_{max}(i)$ (dB)	Band #	Ctr. Freq. (Hz)	$DI_{max}(i)$ (dB)	Band #	Ctr. Freq. (Hz)
/a/ 1% clipping	0.3	2	250	0.4	20	7000
/a/ 8-bit quantization	0.1	15	2900	0.1	13	2150
/u/ 1% clipping	0.3	12	1850	0.3	12	1850
/u/ 9-bit quantization	0.2	1	150	0.2	1	150
/i/ 2% clipping	0.3	13	2150	0.5	9	1170
/i/ 8-bit quantization	1.7	12	1850	1.8	12	1850
3-Vowel Average clipping	0.3	--	----	0.4	--	----
3-Vowel Average quantization	0.7	--	----	0.7	--	----

during the adaptive testing phase, whereas, it was possible to adjust the level of distortion for peak clipping with much greater precision.

It is important to bear in mind that the model has been evaluated, thus far, for single steady-state vowels. Pilot data on other steady-state voiced sounds show similar results, but the critical values of $DI_{max}(i)$ appear to differ for voiceless sounds. Also, the detectability of distortion in single sounds differs significantly from that for running speech; the data for single sounds should be viewed as providing upper or lower bounds on the detectability of distortion in running speech. For example, the data show that 8 bits of quantization are needed for a single steady-state vowel. If, however, 8-bit quantization is provided for running speech then only the most intense sounds will be quantized adequately, while fewer bits of quantization will be available for the less intense sounds. In contrast, if peak clipping is just detectable during the most intense sounds of speech, it will occur less often, or not at all, with less intense sounds and hence be inaudible for all but the most intense sounds.

The data of **Table 2** show that $DI_{max}(i)$ typically occurs in the valley between the formants. The first three formants for the /i/ vowel used in this study were found to occur at frequencies of roughly 390, 2700, and 3900 Hz, respectively (see **Figure 1**). For both forms of distortion, $DI_{max}(i)$ falls in the valley between the first two formants. This is the region in which the adjustment for spread-of-masking effects is greatest (see **Figure 2**). Since the rule for adjusting for spread of masking is based on a simplified approximation of published data, it is of interest to examine the effects of the spread of masking on $DI_{max}(i)$. **Table 2** shows that omitting the spread-of-masking adjustment increases $DI_{max}(i)$ only slightly, but that the frequency region in which $DI_{max}(i)$ occurs does change significantly for the vowel /a/ and to a lesser extent for the vowel /i/.

The general measure of distortion described in this paper is based on a model of the perception of distortion that is similar in form to models proposed for other complex auditory percepts. Examples include models for the prediction of loudness of complex sounds (4,14), the prediction of binaural release from masking and gain in intelligibility for speech in noise (7), the prediction of complex masking effects involving sophisticated signal processing techniques (12), and the detectability of differences in the reverberation characteristics of rooms (10). There also have been several attempts at deriving objective measures of distortion for various speech waveform coders (Levitt (6); Barnwell (1), and references cited therein). It is not coincidental that the proposed measure of distortion draws most heavily on the model of Levitt and Rabiner (7). Further, the similarity between the two models allows for a simple extension of the distortion measure to be used in predicting the loudness of the distortion and the extent to which the distortion interferes with the intelligibility of speech. It is believed that an extension of the distortion index along these lines will provide the framework needed for defining and measuring acceptable levels of distortion above the threshold of audibility.

REFERENCES

1. **Barnwell TP:** Objective measures for speech quality testing. *J Acoust Soc Am* 66:1658–1663, 1979.
2. **Bilger RC and Hirsh IJ:** Masking of tones by bands of noise. *J Acoust Soc Am* 28:623–630, 1956.
3. **Carter NL and Kryter KD:** Masking of pure tones and speech. *J Auditory Res* 2:66–98, 1962.
4. **Florentine M and Zwicker E:** A model of loudness summation applied to noise-induced hearing loss. *Hear Res* 1:121–132, 1979.
5. **Kryter KD:** Methods for the calculation and use of the Articulation Index. *J Acoust Soc Am* 34:1689–1697, 1962.
6. **Levitt H, McGonegal CA, and Cherry LL:** Perception of slope-overload distortion in delta-modulated speech signals. *IEEE Transactions on Audio and Electroacoustics*, AU-18:240–247, 1970.
7. **Levitt H and Rabiner LR:** Predicting binaural gain in intelligibility and release from masking for speech. *J Acoust Soc Am* 42:820–829, 1967.
8. **Penner MJ, Robinson CE, and Green DM:** The critical masking interval. *J Acoust Soc Am* 52:1661–1668, 1972.
9. **Peterson GE and Barney HL:** Control methods used in a study of the vowels. *J Acoust Soc Am* 24:175–184, 1952.
10. **Pierce LK:** *Perception of Coloration in Diotic Reverberant Noise*, Doctoral dissertation, Speech and Hearing Sciences, City University of New York, 1984.
11. **Scharf B:** Critical bands. In *Foundations of Modern Auditory Theory*, Vol I, J. Tobias (Ed.). New York: Academic Press, 1970.
12. **Schroeder MR, Atal BS, and Hall JL:** Optimizing digital speech coders by exploiting masking properties of the human ear. *J Acoust Soc Am* 66:1647–1652, 1979.
13. **Wright HN:** Brief Tone Audiometry. In *Handbook of Clinical Audiology*, 2nd Edition, 218–232, J. Katz (Ed.). Baltimore: Williams and Wilkins Co., 1978.
14. **Zwicker E and Scharf B:** A model of loudness summation. *Psychological Review* 72(1):3–26, 1965.
15. **Zwislocki JJ:** Theory of temporal auditory summation. *J Acoust Soc Am* 32:1046–1060, 1960