Frequency analysis and the critical band*

DAVID R. SODERQUIST
The University of North Carolina at Greensboro, Greensboro, N.C. 27412

The objective of the study was to examine the performance of two groups of Ss on an auditory frequency analysis task. The groups were differentiated in terms of musical training and background. Predictions based on the assumption that frequency analysis is dependent upon the critical band were derived from two synthetically produced inharmonic complex stimuli. Results indicate that nonmusicians are inferior to professional musicians in their ability to analyze complex waveforms. The results suggest that musicians possess critical bands which are rectangular in shape and approximately 20% narrower in width than published values.

In 1843, Ohm presented his acoustic law stating that the auditory system may be thought of as a Fourier analysis device. This law simply says that the ear is assumed to have the ability to analyze and separate components of a periodic complex stimulus and identify each component. This ability, commonly known as frequency analysis (FA), has led to considerable discussion and debate concerning the extent to which the auditory system can analyze complex stimuli. Carried to an extreme, Ohm's law states that the ear should accomplish an impossible task, viz, detect each component frequency in a multifrequency periodic waveform. Obviously the analysis capability of the ear is limited. Plomp (1964) examined the limits and suggested that FA may be dependent upon a functional parameter of the auditory system called the critical band (CB).

Plomp's suggestion leads to a logical conclusion that performance on a FA task is predictable. All that is required is that: (1) the component frequencies that comprise the complex waveform be stipulated and (2) the CB associated with each component frequency in the complex waveform be known. Given these two parameters, a prediction may be derived concerning the number of components in a complex waveform that should be detectable, regardless of whether the complex stimulus is periodic or not. Consequently, two complex waveforms were produced, and predictions were made concerning the detectability of individual components comprising the stimuli. Tables 1 and 2 show the complex stimuli and illustrate how the predictions were derived. The tables show the 12 component frequencies that were combined to form

complex waveforms and the respective CB for each component as estimated by Zwicker (1961). The prediction assumes that each CB integrates energy over its bandwidth, and a decrease in detectability of a component frequency occurs if its associated CB overlaps with another CB.[1] If CBs do not substantially overlap, then individual component frequencies are assumed to be detectable. According to this rationale, it would be predicted that a listener should detect about the first five components of the complex waveform. In contrast, frequencies of the complex above the fifth component should be masked, and performance on the analysis task should deteriorate to chance.[2]

Prior to testing the prediction, S population was recognized as being an important consideration. This was particularly true in view of Plomp's suggestion that FA is dependent on the CB. Review of literature reveals, for example, that musical background or aptitudes of Ss used in research to deliniate CBs were seldom, if ever, reported (cf. Scharf, 1961; Zwicker, Flottorp, & Stevens, 1957). Hence, the major objective of the experiment was to examine empirically FA ability of two groups of Ss differentiated in terms of musical background and training and to test predictions derived from two synthetically produced inharmonic complex stimuli.

METHOD

Two groups, each consisting of four male Ss, were paid listeners. Group 1 was naive (not musically trained) and ranged between the ages of 22 and 27. Group 2 consisted of professional musicians between the ages of 23 and 31. Each musician had been playing musical instruments for an average of 18 years.

Twelve component complex stimuli were produced by using six Hewlett-Packard Model 201-C tone generators in conjunction with a two-channel Ampex Model AG-3520 magnetic tape recorder. Each channel of the recorder was used to record 6 of the 12 component frequencies. The outputs of the two channels were mixed and fed through a Krohn-Hite Model 3100 bandpass filter to a calibrated earphone mounted in a MX-41/AR cushion. The frequencies and amplitudes of the components were monitored by Hewlett-Packard Model 5223-L electronic counter and a Ballentine Model 300-H true RMS voltmeter. The component frequencies were all approximately 50 phons. Figure 1 shows a block diagram of the experimental equipment.

The two sets of complex stimuli produced by the above procedure are shown in Tables 1 and 2. The component frequencies were deliberately chosen so the prediction from the CB concept would be

TABLE 1
Complex Stimulus A

Component Number	Component Frequency	Estimated Critical Bandwidth
1	240	100
2	438	108
3	720	141
4	920	155
5	1120	184
6	1310	204
7	1480	224
8	1650	248
9	1800	272
10	1960	295
11	2065	308
12	2200	328

TABLE 2
Complex Stimulus B

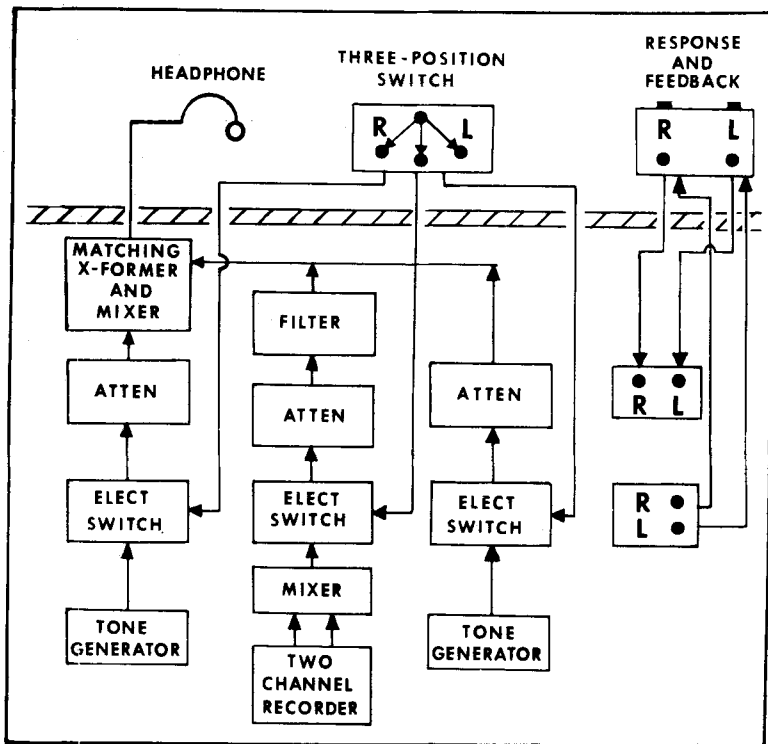| Component Number	Component Frequency	Estimated Critical Bandwidth
1	225	100
2	410	106
3	708	140
4	880	151
5	1100	182
6	1295	203
7	1465	223
8	1640	244
9	1785	268
10	1945	290
11	2050	300
12	2185	327

Fig. 1. Block diagram of experimental apparatus.

the same for each complex stimulus. In addition, the components comprising the complex waveforms were purposely inharmonic.

Plomp's (1964) procedure was used as a methodological model. Each S was seated in a sound-shielded room and had at his disposal a three-position stimulus presentation switch and response buttons. The center position of the switch keyed a Grason-Stadler Model 829-C electronic switch allowing the synthetic complex stimulus to be heard by S; the left and right positions keyed independent electronic switches, allowing S to hear one of two different pure tone comparisons. The separate presentation of comparison tones and complex stimuli thus eliminated the opportunity to detect components in the complex via beats; the use of an inharmonic complex eliminated the possibility of S using harmonic ratios as an aid in "hearing out" the component frequencies. The two-alternative forced-choice procedure required S to decide which of two comparison signals coincided in frequency with one of the components of the complex stimulus. The dependent variable was percentage correct, and no time limits were imposed. S indicated his decision by depressing a response button corresponding to his choice, and immediate feedback was provided. The correct comparison was randomized left and right with a probability of .5 and also randomized so

that higher comparison frequencies were correct on half of the trials.

Because FA is a difficult task, training was used prior to the experiment itself to acquaint Ss with the task. A total of 308 training trials, ranging from easy (36 trials) to moderately difficult (96 trials) to difficult (176 trials), were given each S.

The difficulty of the training stimuli was determined by the number and frequency spacing of the components in three different synthetically produced complex stimuli. The component frequencies of the training stimuli were different from those of the experiment proper. The experiment itself consisted of 240 trials for each S, 20 comparison trials for each of the 12 component frequencies in the complex, with frequent rest periods to avoid fatigue.

The experiment was completely replicated with the two complex stimuli. Half the Ss in each group received Complex Stimulus A and the other half received Complex Stimulus B.

At the completion of the experiment, a short experimental test was given to examine the possibility that Ss had learned some of the correct choices as a result of repeated presentations with feedback. The test was brief and consisted of 48 trials (four for each component in the complex) without the complex being presented and without feedback. The S was simply instructed to state which of the two comparison tones on each trial would be correct if the complex were in fact presented.

RESULTS

A repeated-measures analysis (Winer, 1962, p. 337) revealed no significant differences between groups or between complex stimuli; there were no significant interactions. However, there was a significant difference for component frequencies (p < .01). Based on this
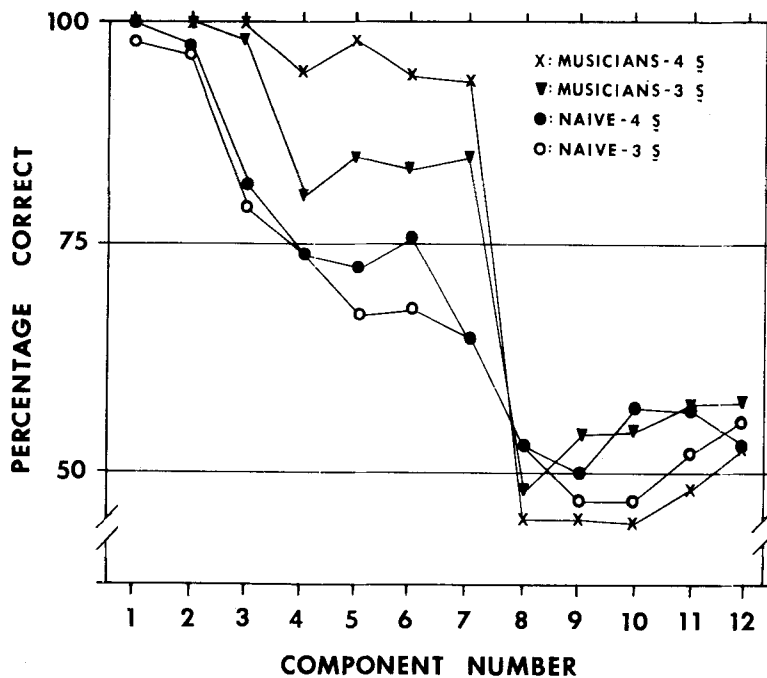

Fig. 2. Percentage correct as a function of component number in the inharmonic complex stimuli.

analysis and an examination of individual performance, the data for each group were averaged over the two stimuli and plotted graphically as a function of the component number. Figure 2 shows that, when using a criterion of 75% correct, four naive Ss detected approximately five components, whereas the four musicians detected approximately 7 of the 12 components in the complex. However, examination of the data in terms of the postexperimental test revealed that two Ss, one in each group, may have learned some of the components as a result of repeated exposure with feedback.[3] The data were reanalyzed with these two Ss eliminated. The analysis (Winer, 1962, p. 302) revealed significant differences among the partials ($p < .01$), and marginally significant differences between groups ($.05 < p < .10$). In addition, a significant interaction ($p < .05$) led to a series of simple main effects tests (Winer, 1962, p. 311) which revealed that Components 3 through 7 were significantly different ($p < .01$) and the simple effects for components at the two levels (groups) were significant ($p < .01$). The meaning of this second analysis is also shown in Fig. 2. The figure substantiates the analyses by showing the clear difference between the two groups (three Ss per group) in terms of the number of detectable components. The interaction is also shown in that a S's performance in detecting Components 3 through 7 is dependent on group membership. In general, the data indicate that musicians detected about seven components and the naive about four. Removal of one S from each group resulted in less variability within each group and no substantial change in the results.

DISCUSSION

The agreement between the prediction and the data for naive Ss is not precise. The data reveal that the actual number of components detected is less than five and closer to four. Whether this deviation from prediction is large enough to warrant the conclusion that frequency analysis is not dependent on the CB is problematical. However, in view of the variability in the reported widths of CBs (cf. Swets, 1963; Swets, Green, & Tanner, 1962) and the strong possibility that CBs are not rectangular (cf. Greenwood, 1961; Scharf, 1961), the deviation does not appear to justify the rejection of the conclusion that frequency analysis is dependent on the CB. Thus, Plomp's results have been replicated.

The conclusion that musicians are superior to naive Ss on a FA task appears to be a reasonable conclusion. The reason for the superiority could, of course, be attributed to either musical training (perceptual learning) or innate ability. Although the answer to this question is not provided by this experiment, it is of interest to consider the musician's performance in relation to the CB concept.

If one accepts the postulate that the CB determines the limits of FA for both musicians as well as nonmusicians, then a logical conclusion is that CBs for musicians are both narrower and more rectangular than those of naive individuals. Using this logic and a post hoc inspection of musicians' performance, the original prediction may be altered to fit the data by assuming musicians' CBs are rectangular in shape and approximately 20% smaller than the published values (Zwicker, 1961). This assumption leads to a prediction which fits the obtained data for musicians quite well. It not only accounts for the reason why more components are detected, but also for the high percentage correct on the first seven components and the abrupt drop to chance performance thereafter.

REFERENCES

FLETCHER, H. Auditory patterns. Review of Modern Physics, 1940, 12, 47-65.

GARNER, W. R. The effect of frequency spectrum on temporal integration of energy in the ear. Journal of the Acoustical Society of America, 1947, 19, 808-815.

GÄSSLER, G. Ueber die Horschwelle für Schallereignisse mit verschieden breitem Frequenzspektrum. Acustica, 1954, 4, 408-414. Cited by B. Scharf, Complex sounds and critical bands. Psychological Bulletin, 1961, 58, 205-217.

GREEN, D. M., BIRDSALL, T. G., & TANNER, W. P., JR. Signal detection as a function of signal intensity and duration. Journal of the Acoustical Society of America, 1957, 29, 523-531.

GREENWOOD, D. D. Auditory masking and the critical band. Journal of the Acoustical Society of America, 1961, 33, 484-502.

PLOMP, R. The ear as a frequency analyzer. Journal of the Acoustical Society of America, 1964, 36, 1628-1636.

PLOMP, R., & BOUMAN, M. A. Relation between hearing threshold and duration for tone pulses. Journal of the Acoustical Society of America, 1959, 31, 749-758.

SCHARF, B. Complex sounds and critical bands. Psychological Bulletin, 1961, 58, 205-217.

SHEELEY, E. C., & BILGER, R. C. Temporal integration as a function of frequency. Journal of the Acoustical Society of America, 1964, 36, 1850-1857.

SWETS, J. A. Central factors in auditory frequency selectivity. Psychological Bulletin, 1963, 60, 429-440.

SWETS, J. A., GREEN, D. M., & TANNER, W. P. On the width of critical bands. Journal of the Acoustical Society of America, 1962, 34, 108-113.

WINER, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

ZWICKER, E. Subdivision of the audible frequency range into critical bands. Journal of the Acoustical Society of America, 1961, 33, 248.

ZWICKER, E., FLOTTORP, G., & STEVENS, S. S. Critical band width in loudness summation. Journal of the Acoustical Society of America, 1957, 29, 548-557.

NOTES

1. The assumption that the CB integrates energy over its bandwidth, although not a strict necessity for the prediction, appears to be parsimonious in view of Fletcher's (1940) original conception. In addition, there are supporting data for such an integration hypothesis (Garner, 1947; Gässler, 1954; Green, Birdsall, & Tanner, 1957; Plomp & Bouman, 1959; Sheeley & Bilger, 1964).

2. The prediction was based on the assumption that critical bands are rectangular in shape. This assumption may possibly be untenable in view of available data (Greenwood, 1961; Scharf, 1961). However, it is a useful postulate in making an initial performance prediction.

3. The probability of a S getting four correct out of four presentations by chance is .0625. Thus, if a success is defined as four correct out of four presentations, a binomial distribution may be determined and the probability of obtaining x successes out of N possibilities calculated. The exact probability may then be used as a guide to determine whether S performed above chance on the postexperimental test. The two Ss eliminated both had three or more successes as defined above. The probability of such a result by chance is less than .04. None of the remaining Ss received more than two successes. (Probability of getting two or fewer successes is, of course, $1.00 - .04 = .96$.)