The MIT Press

The Fusion and Layering of Noise and Tone: Implications for Timbre in African Instruments
Author(s): Cornelia Fales and Stephen McAdams
Source: *Leonardo Music Journal*, Vol. 4 (1994), pp. 69-77
Published by: The MIT Press
Stable URL: http://www.jstor.org/stable/1513183
Accessed: 03/06/2010 10:44

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/action/showPublisher?publisherCode=mitpress.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

The MIT Press is collaborating with JSTOR to digitize, preserve and extend access to *Leonardo Music Journal.*

STOR®
http://www.jstor.org

SOUNDING THE MIND

The Fusion and Layering of Noise and Tone: Implications for Timbre in African Instruments

Cornelia Fales
and Stephen McAdams

ABSTRACT

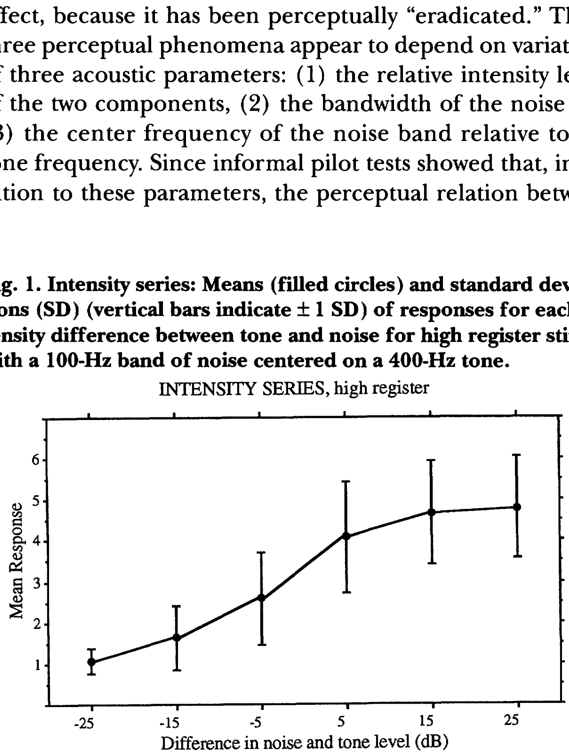
Since their earliest explorations of African music, Western researchers have noted a fascination on the part of traditional musicians for noise as a timbral element. The authors present the results of perceptual and acoustic investigations of the fusion and "layering" of noise and tone. These results have implications for pitch and timbre in both traditional and non-traditional, acoustic and synthesized music. The results define possible perceptual relations between noise and tone and reveal that the construction of noise devices should follow relatively precise acoustic rules involving the frequency, the bandwidth and the level of the noise relative to those of the tone. The results also exemplify the fusion of two extremely different timbres, with implications for the blending of instrumental timbres in an orchestral setting. The experiments should be of interest to composers who synthesize mixtures of noise and periodic sound and for whom the control of such mixtures remains problematic.

From their first contact with African music, Western ethnomusicologists have remarked on the predilection of African musicians to "layer" (or superimpose) musical and non-musical sound until the distinction between them is lost [1]. In particular, it has been observed that African music shows a special fascination with "noise"—the transformation of ordinary, mundane sounds into the substance of music. And while the manufacture of classical instruments and the performance practice of Western musicians has aimed toward reducing the amount of extraneous noise produced by an instrument, African musicians augment the natural noise potential of their instrument by attaching noise-makers such as rattling seeds or bottle caps on which the vibrations of the main resonator operate. The effect is a complicated layering of sound, rich in aperiodic complexity.

One of the motivations for our study was the realization that the combination of noise and musical elements, traditionally described by ethnomusicologists as "layering," actually takes at least two perceptual forms. Either the sound is truly layered, and listeners hear two or more perceptually distinct sounds concurrently, or the physically superimposed sounds are perceptually fused, so that listeners hear a single sound—a blend of the two sounds—neither of which is identifiable as the primary or the superimposed sound. A more commonly recognized example of this distinction is the difference between a voice speaking through or in the presence of noise (layered noise and voice) and a hoarse voice speaking (noisy voice).

The direction of our research has involved both perceptual experiment and acoustic analyses. Like many psychoacoustic studies, the first (experimental) part of our investigation involved the use of artificial stimuli, reduced to controllable acoustic variables. To complement this part, an acoustic analysis of real African instruments in which noise plays a prominent part was essential to verify the relevance of our experimental results to sounds actually produced by those instruments. In particular, our experiments were geared toward establishing the parameters of noise and tone that influence the perceptual relation between them. We considered three possible relations between noise and tone: (1) *fusion* (the noise and tone are perceptually integrated into a single sound event, both contributing to the perceptual quality of the event), (2) *layering* (the two components are perceptually segregated into distinct percepts, each with its own perceptual qualities) and (3) *masking* of tone by noise (the more intense noise "covers up" the tone that can thus no longer be heard *and* has no perceptual effect on the noise component). Layering is distinguished from fusion and masking in that a layered tone can be heard separately from the noise. While a tone that is fused with a noise cannot be heard separately, it still contributes to the timbral quality of the noise; a masked tone has no such effect, because it has been perceptually "eradicated." These three perceptual phenomena appear to depend on variations of three acoustic parameters: (1) the relative intensity levels of the two components, (2) the bandwidth of the noise and (3) the center frequency of the noise band relative to the tone frequency. Since informal pilot tests showed that, in addition to these parameters, the perceptual relation between

Fig. 1. Intensity series: Means (filled circles) and standard deviations (SD) (vertical bars indicate ± 1 SD) of responses for each intensity difference between tone and noise for high register stimuli with a 100-Hz band of noise centered on a 400-Hz tone.



Cornelia Fales (ethnomusicologist), Institut de Recherche et Coordination Acoustique/Musique (IRCAM), 1 place Stravinsky, F-75004 Paris, France.

Stephen McAdams (experimental psychologist), Laboratoire de Psychologie Expérimentale (CNRS), Université René Descartes, EPHE, 28 rue Serpente, F-75006 Paris, France.

Manuscript solicited by Kathryn Vaughn.

Received 11 May 1994.

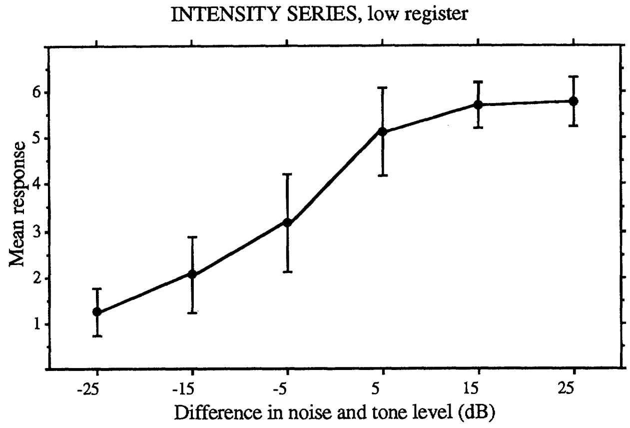


Fig. 2. Intensity series: Means (filled circles) and standard deviations (SD) (vertical bars indicate ± 1 SD) of responses for each intensity difference between tone and noise for low register stimuli with a 100-Hz band of noise centered on a 1,000-Hz tone.

noise and tone was greatly affected by the pitch register of the two elements, each of the three parameters were varied for stimuli constructed around both high and low registers (centered at 1,000 Hz and 400 Hz, respectively).

PERCEPTUAL RESEARCH

General Method

The experiments were organized in three series, each testing the effect of a different acoustic parameter on tone/noise fusion. In the Intensity Series, pure tones (sinusoidal waveform) in high and low registers were combined with band-filtered white noise having a bandwidth of 100 Hz and a center frequency coinciding with the frequency of the tone. For this series, the intensity level of the tone varied randomly within a given range, while the noise level was offset from the tone level by specific increments. For the Bandwidth Series, this procedure was repeated for several noise bandwidths ranging from 50 to 200 Hz. For the Center Frequency Series—in which the noise center frequency was tested relative to the tone frequency—the tone varied as described above in high and low registers, and the center frequency of the noise was offset from the tone by up to 200 Hz in the low register and up to 400 for the high register. The noise in this last series had a fixed bandwidth of 100 Hz.

On a given trial in one of the series, a tone/noise stimulus was presented three times in succession. Each stimulus was 900 msec in duration and the onsets and offsets of the noise and tone components were synchronous. The tone was presented at a mean level of 57 dB SPL (varying randomly by ± 9 dB around this value on a trial-to-trial basis), and the noise level was determined with respect to this level for each trial.

The perceptual research was performed at Institut de Recherche et Coordination Acoustique/Musique (IRCAM), a music research institute in Paris. We therefore were initially constrained in the listeners we could readily test, being limited primarily to Europeans and North Americans. Since the aim of this part of the work was to establish some basic psychoacoustic thresholds and because psychoacoustic phenomena at this level are generally thought to be a function of the human hearing apparatus (which is essentially the same for all human beings), we considered the results obtained from this study to reflect universal properties of auditory perception—although this assumption warrants verification in future research. From an ethnomusicological point of view, however, it is noteworthy that informal pilot tests with these subjects revealed that the notion of fusion between dissimilar timbres was a difficult concept to grasp intellectually and, even more so, perceptually. Initially, subjects were asked to rate on a continuum the "degree of fusion" demonstrated by a series of stimuli, but it became evident that the subjects were unclear as to the definition of the perceptual phenomenon that they were asked to judge. Therefore, as an alternative, the subjects were asked whether the acoustic components comprised one or two sounds, since fusion transforms multiple elements into a single unitary percept. Subjects were asked, in other words, to rate the degree to which the tone could be heard separately from the noise in each trial.

Twenty-three listeners (all but one had had some musical training; none were professional musicians) were instructed to decide whether they heard the tone separately from the noise or not and to rate their certainty by selecting one of six buttons as follows: (1) "Tone heard separately: I am sure"; (2) "Tone heard separately: I am fairly sure"; (3) "Tone heard separately: I am not sure"; (4) "Tone NOT heard sepa-

Fig. 3. Bandwidth series: Means (filled circles) for response categories 3 and 4 (indicating fusion) for each bandwidth as a function of level difference between tone and noise for high register. Vertical axis shows amplitude difference between tone and noise; horizontal axis indicates bandwidth.

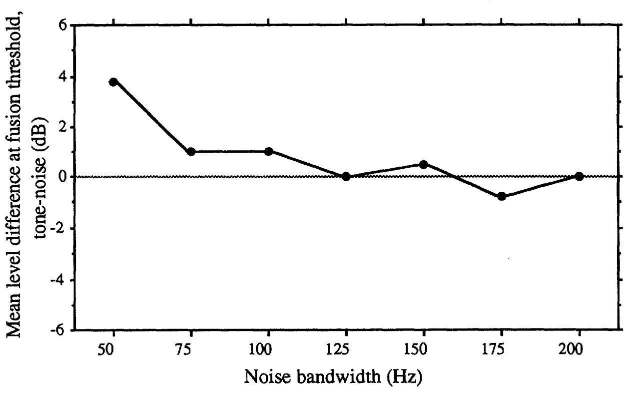
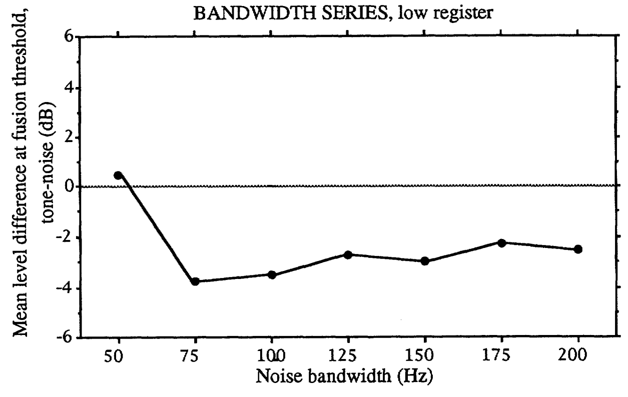


Fig. 4. Bandwidth series: Means (filled circles) for response categories 3 and 4 (indicating fusion) for each bandwidth as a function of level difference between tone and noise for low register. Vertical axis shows amplitude difference between tone and noise; horizontal axis indicates bandwidth.



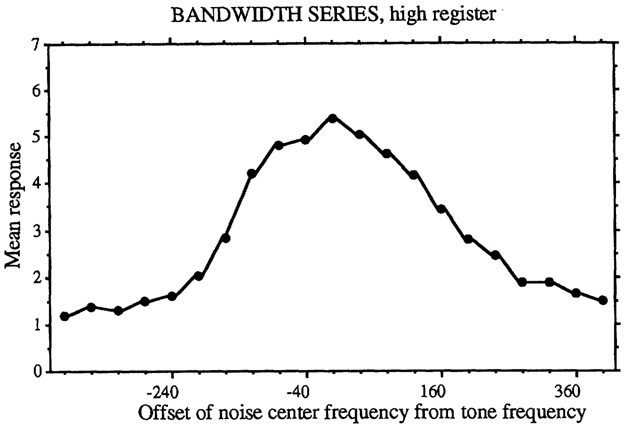


Fig. 5. Center Frequency Series: Means (filled circles) of responses for high register stimuli in which center frequency of noise bandwidth is offset from tone frequency by up to 400 Hz. Vertical axis shows response, horizontal axis shows offset of noise-center frequency from tone.

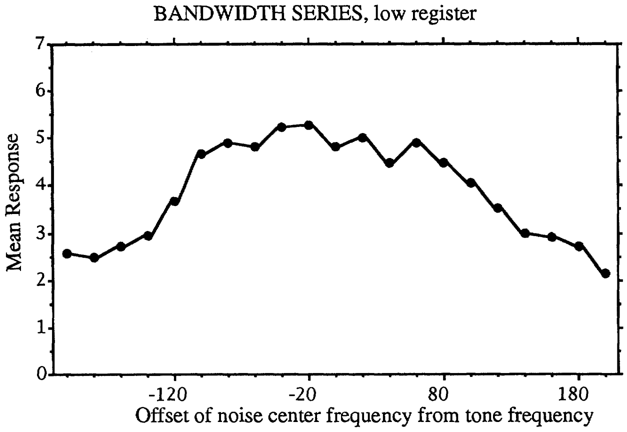


Fig. 6. Center Frequency Series: Means (filled circles) of responses for low register stimuli in which center frequency of noise bandwidth is offset from tone frequency by up to 200 Hz. Vertical axis shows response, horizontal axis shows offset of noise-center frequency from tone.

rately: I am not sure"; (5) "Tone NOT heard separately: I am fairly sure"; (6) "Tone NOT heard separately: I am sure." We call these judgments "tone separation ratings." For the final Center Frequency Series, listeners were alerted, in addition, that the range of perceptual differences might be changed, since there might no longer be a condition in which the tone was not heard at all, as occurred in the first two series of trials.

Intensity Series

Tone separation ratings were made for 11 intensity differences between tone and noise varying from –25 to +25 dB in 5-dB increments. Each condition was repeated five times for each listener. These responses were then averaged across listeners and repetitions. The mean responses and their standard deviations (a measure of variation across responses for a given condition) for each intensity difference are shown in Figs 1 and 2 for high and low registers, respectively. Note that when the noise is weak compared to the level of the tone, low ratings are given (tone heard separately) and when the noise is intense, high ratings are given (tone not heard separately). The curves for both registers suggest a gradual progression of the percept from separation through fusion to masking.

Presuming that high ratings (5–6) indicate that the tone may be masked by the noise and that low ratings (1–2) indicate that the tone is heard separately from the noise, we take intermediate ratings (3–4) as indicative of the conditions under which the tone and noise are beginning to fuse. Figures 1 and 2 show that the degree of noise-tone relation in the high and low registers differs most dramatically on the high intensity-difference end of the curve. For the low register, when the noise exceeds the tone by more than about 4 dB, the tone begins to lose audibility; when the noise intensity is less than the tone by about 15 dB, the tone separates and the sound is layered. Based on intermediate ratings, a first estimation of the necessary intensity relation between noise and tone, therefore, puts the fusion region between –7 and –1 dB, noise relative to tone. For the high register, fusion appears to occur when the intensity level of the noise is between –3 and +4 dB, relative to the tone. There would appear, therefore, to be a difference between high and low registers, fusion occurring at a slightly higher intensity difference (about 3 dB) for the high register than for the low register.

Bandwidth Series

Tone separation ratings were made for 11 intensity differences as in the previous series, but here they were performed for each of five noise bandwidths varying from 50 to 200 Hz in 50-Hz increments. The noise bands were created with a second-order bandpass filter. Each condition was repeated three times for each listener. These responses were then averaged across listeners and repetitions. For this series, the same types of mean response curves resulted, with the same difference in registers, as were seen for the first series with a noise bandwidth of 100 Hz. In order to compare the intensity differences between noise and tone in the fusion region (responses 3–4) across the different bandwidths, the intensity differences corresponding to mean responses of 3 and 4 were averaged across subjects and plotted in Figs 3 and 4 for the low and high registers, respectively. This comparison shows that the difference in level between tone and noise at the fusion threshold decreases slightly for both the high and low registers as the bandwidth increases from 50 to 75 and then remains relatively constant for larger bandwidths.

Center Frequency Series

In this experimental series, a constant bandwidth (100 Hz) and noise-intensity level (65 dB with respect to the tone level) were maintained, while varying the difference between the center frequency of the noise and the tone frequency from –200 to +200 Hz in increments of 20 Hz for the low register and from –400 to +400 Hz in increments of 40 Hz for the high register. It is important to note that for the stimuli in this series, there was no condition in which the noise masked the tone, since the intensity level of the noise remained constant at a level difference found to be within the fusion response area for a bandwidth of 100 Hz in the first experimental series. Thus, for this series, response category 6 (Tone NOT heard separately: I am sure) was taken to indicate the greatest degree of fusion, and response category 5 (Tone NOT heard separately: I am fairly sure) was considered the fusion threshold. In this series, in other words, judgments of non-separation were taken to indicate fusion rather than masking, given the signal parameters used.

Mean responses across listeners and repetitions as a function of the frequency difference are shown in Figs 5

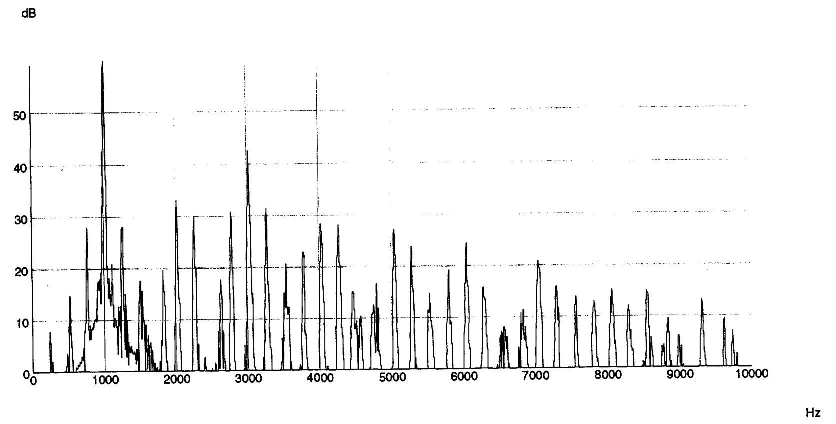


Fig. 7. Narrowband spectrum of *harmonic* elements of flute tone of fundamental frequency 1,032 Hz. The horizontal axis corresponds to frequency (in Hz), and the vertical axis to amplitude (in dB). The fundamental frequency is approximately 1,032 Hz.

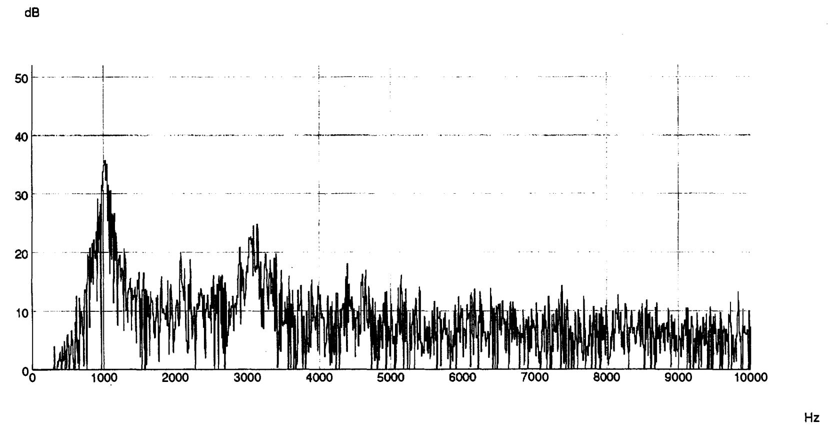


Fig. 8. Narrowband spectrum of *noise* elements of flute tone of fundamental frequency 1,032 Hz. The horizontal axis corresponds to frequency (in Hz), and the vertical axis corresponds to amplitude (in dB). Note the similarity in envelope (outline of peaks) of the noise spectrum to the harmonic spectrum above, indicating that noisebands are roughly centered on primary harmonics.

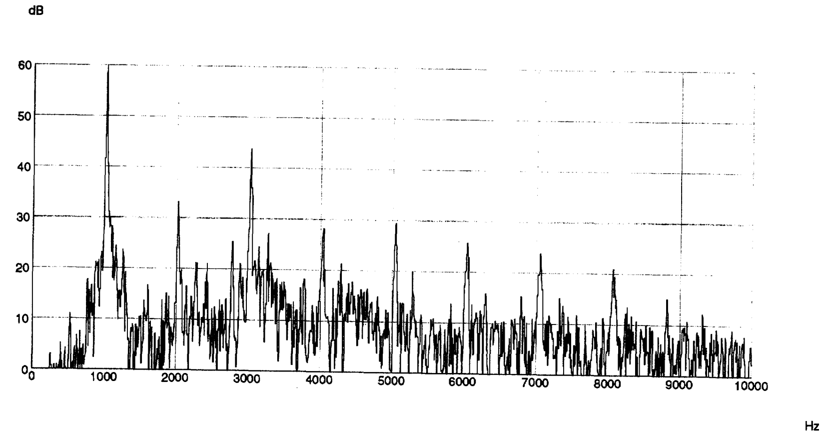


Fig. 9. Narrowband spectrum of flute tone (*complete*) of fundamental frequency 1,032 Hz. The horizontal axis corresponds to frequency (in Hz), and the vertical axis to amplitude (in dB). Note presence of a second set of "harmonics," which are marked by arrows. Note also that the width of noisebands surrounding fundamental and second two harmonics are roughly as large as the bandwidths of the auditory filters centered on these components.

and 6 for both registers. These curves demonstrate that fusion is affected by the way the auditory system filters the incoming sound into frequency bands in a manner similar to that shown for masking by a vast body of psychoacoustic research. The peripheral auditory system can be considered to be a bank of overlapping bandpass filters (called "auditory filters") whose center frequencies cover the audible frequency range and whose bandwidths (often called "critical bandwidths") vary systematically with the center frequency, being smaller at low frequencies and larger at high frequencies. Thus, as the center frequency of the noise moves away from the tone frequency and its energy starts to fall outside the auditory filter center on the tone, the degree of fusion decreases. This effect gives rise to the bell-shaped curves in Figs 5 and 6.

Discussion of the Perceptual Experiments

From an ethnomusicological point of view, a primary interest of the results obtained so far resides in the relatively small range of the acoustic parameters that determine fusion. While we have not yet tested the three parameters as they might operate for fusion with complex tones, we can hypothesize that if a musician were to construct an ideal noise device intended to fuse with a continuum of *pure* tones that made up the pitch inventory (an altogether unrealistic scenario), the device would produce noise having the following characteristics: (1) a bandwidth as large or larger than that of an auditory filter centered on the highest pitch of the instrument; (2) a center frequency as close as possible to the frequency of the tone with which it is to fuse, but with a frequency offset small enough to still allow the required noise power to enter an auditory filter centered on the lowest frequency produced by the instrument; and (3) an intensity level that allows the required noise power to enter the auditory filter centered on the highest frequency produced by the instrument, given the chosen bandwidth of the noise. Informal testing of laboratory stimuli has demonstrated that the fusion of noise with a more realistic complex tone is more complicated than fusion with a pure tone. It is not a simple matter of surrounding each harmonic with noise or adding a single block of noise with bandwidth equal to the bandwidth of the complex tone. However, though confirmation will only be possible with continued experimentation, we suggest that to the extent that the noise phenomena of fusion and masking are related, then our results on the fusion of a pure tone might be relevant to complex tones as well. Since there are indeed instruments (such as the sanza of Southern, Central, and West Africa, an example of which we will discuss below) in which a unitary noise device contributes a cross-frequency "buzz" to the timbre, it is probable that the invention of such devices is neither haphazard nor epiphenomenal. It does not appear to be the case that the use of fused noise, at any rate, is simply another example of the African predilection for layered textures. Rather, as demonstrated below, such devices must be carefully chosen or constructed in order to achieve a desired effect.

ACOUSTIC ANALYSES

The second part of our study aims to complement the perceptual research with acoustic analyses of noise devices in African musical instruments. The analyses whose results are presented here were done using two kinds of digital signal processing software developed at IRCAM. The first of these (which we will call "noise-separation") separates the non-periodic part of the signal (the noise) from the periodic part (harmonic tone). The second is a filtering program that allows very fine pass- or stopband filtering by literally "drawing in" the desired filter on a spectrum. We present the results from three instruments, a traditional bamboo flute, a sanza [2] with a single level of metal *lamellae* (tongues) and a musical bow [3]. All three intruments are from Burundi and were chosen for analysis because the first two demonstrate fused noise and tone, and the last demonstrates layered noise and tone.

Regarding the fused instruments, the flute has a relatively simple spectrum, while the sanza is more complicated. Perceptually, the flute timbre is characterized by breathiness or air noise, which—although weaker in regard to fusion than the noise in our synthesized examples—demonstrates several parameters of concern in our fusion experiments. The sanza noise is produced by beer bottle caps that have been attached to the resonator of the instrument. Noise-tone fusion in the case of this instrument is perceptually much stronger than for the flute. For the musical bow, which demonstrates unfused noise and tone and whose spectrum is also relatively complicated, noise is produced by metal "clackers" attached to the gourd resonator, producing a metallic jingling sound superimposed over the sound of the struck cord.

Ironically, the major problem encountered in noise analysis of real instruments is exactly the result of the phenomenon that is the object of our research: both acoustically and perceptually, noise is intertwined with other elements of the music with tremendous complexity. Whether the noise is layered or fused, the resulting acoustic signal, of course, includes both noise and tonal components, and the distinction between them is difficult to detect. The noise-separation software that we used operates on the assumption that the constituent noise elements are lower in amplitude, within a given frequency region, than the harmonic elements. The results of our perceptual experiments show that it is possible to achieve perceptual fusion of noise and tone even when the noise is at a higher level than the tone. Indeed the last instrument we will discuss—the musical bow—shows noise elements that are more intense than periodic elements in some frequency regions. In the case of the musical bow, then, the noise-separation software was unable to distinguish periodic from aperiodic elements, and we used filtering software to isolate noise from the harmonic elements. Even in the case of the flute and sanza, where the noise separation program was used, the samples had to be resubmitted for processing several times, in order to separate noisy from periodic elements as much as possible.

We will look first at the instruments in which noise and tone fuse. Figures 7–12 below show spectra of the flute and sanza. In the first group of three, Fig. 7 shows the harmonic elements only of the flute tone with a fundamental frequency of 1,032 Hz; Fig. 8 shows the aperiodic content of the same tone; Fig. 9 shows the entire conglomerate. The second three figures show, respectively, the harmonic structure (Fig. 10), the aperiodic content (Fig. 11) and the entire sanza tone (Fig. 12) with a fundamental frequency of 285 Hz. The figures for both instruments show that primary harmonics are surrounded by noise. Both noise spectra, in fact, appear to follow the harmonic spectra of their respective instruments, so that the envelope of just the noise portion of each signal resembles the envelope of the related harmonic portion, though at a lower amplitude.

Flute Analyses

For the flute, the width of the noise band surrounding the fundamental and the second two harmonics is at least as large as the bandwidths of the auditory filters centered on these components (for the fundamental frequency at 1,032 Hz, the noise bandwidth must be approximately 136 Hz; for the second [4] harmonic at 2,016 Hz, the bandwidth must be approximately 241 Hz; for the third harmonic at 3,105 Hz, the bandwidth must be 349 Hz). In addition, the noise in the frequency region of the fundamental and third harmonic has the highest level, corresponding to the fact that these are the strongest harmonics. Therefore, for the perception of the entire complex conglomerate, the auditory filters in the region of the fundamental and the third harmonic are the most stimulated by the noise. In the case of the higher harmonics, it is difficult to determine where the noise bandwidth begins, since the noise level flattens out considerably above about 5,000 Hz.

We can make several other observations based on the spectra, as well as on the sound of the signal we are analyzing. As mentioned above, the noise element in the timbre of this flute is slight, as might be predicted from the level of the noise. If the fusion of noise with a complex tone has some relation to noise–pure tone fusion, then, according to our experiments, the difference in noise and tone level here ought to make this combination difficult to fuse. And yet, they *are* fusing, as is evidenced by the sound of the entire signal as well as the sound of the separate noise and tone signals (both of which are noticeably altered by their isolation from one another). From this example, we can arrive at one of two conclusions. It may be that noise fusion in the case of a complex tone differs markedly from noise fusion with a pure tone. However, insofar as the noise phenomena of fusion and masking are related, Moore's observation [5] that the masking of a complex tone by noise is predictable from the detectability of the most prominent harmonics may be equally applicable to fusion. Therefore, our results on the fusion of a pure tone *ought* to be relevant to complex tones as well.

An alternative explanation of the failure of the flute example to conform more exactly to our experimental results is that these results may need to be rein-

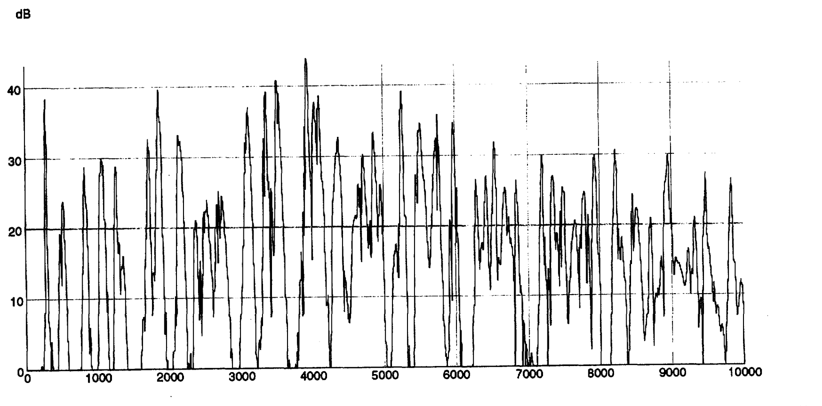


Fig. 10. Narrowband spectrum of *harmonic* elements of sanza tone of fundamental frequency 285 Hz. The horizontal axis corresponds to frequency (in Hz), and the vertical axis to amplitude (in dB). Note the relative inharmonicity of partials, as indicated by their uneven distribution across the frequency axis.

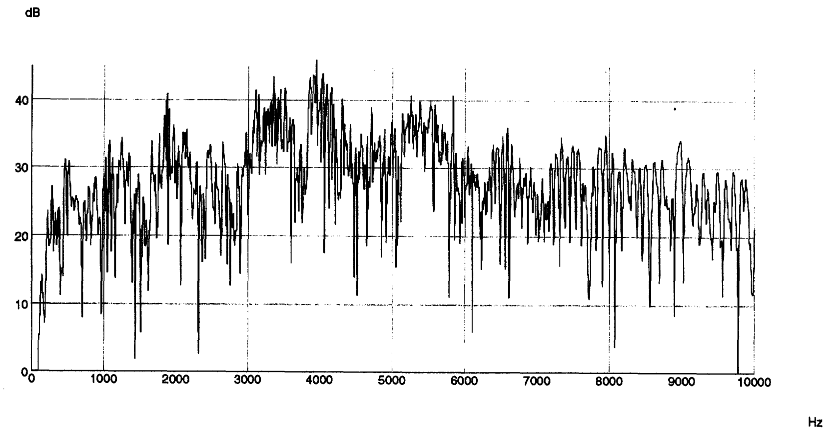


Fig. 11. Narrowband spectrum of M *noise* elements of sanza tone of fundamental frequency 285 Hz. The horizontal axis corresponds to frequency (in Hz), and the vertical axis to amplitude (in dB). Notice the similarity in envelope (outline of peaks) of the noise spectrum to the harmonic spectrum above, indicating that noisebands are roughly centered on primary harmonics.

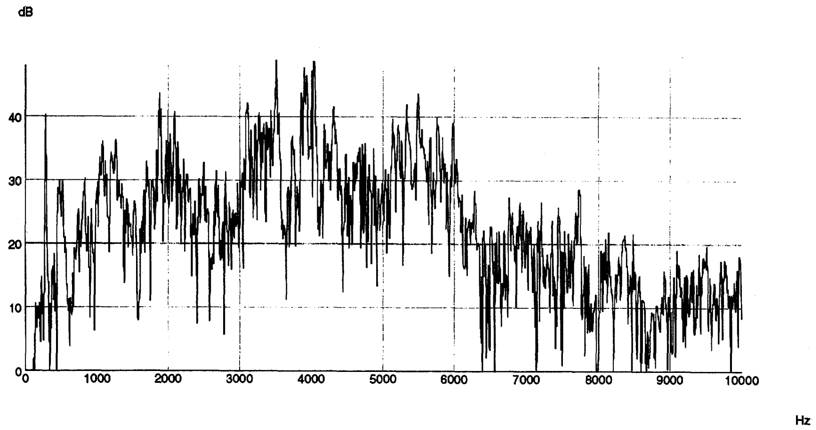


Fig. 12. Narrowband spectrum of sanza tone (*complete*) of fundamental frequency 285 Hz. The horizontal axis corresponds to frequency (in Hz), and the vertical axis to amplitude (in dB). Here noise bandwidths begin to approximate the size of respective auditory filters only starting at the fifth harmonic.

terpreted to account for *degrees* of fusion. In order to do this, we need to establish a definition of fusion that accounts for the varying strengths of listeners' sensations. Does "weak fusion" occur when the noise accompanying a periodic tone is only lightly influential in the perceptual composition of the timbre? And does that mean that the noise itself must be at a low level? Or can the noise be relatively intense, but with its center frequency offset by such a degree from the tone that only a part of it is contributing to the sensation of fusion? If the second arrangement of noise and tone is allowed as a kind of fusion with only part of the noise fusing, then can noise both fuse and layer with the tone it accompanies? In our experiments, we have considered all stimuli provoking responses 3 or 4 to be fused, but perhaps response 2 ought to be considered to reflect a less strong degree of fusion—thus provoking a less sure sensation of the separateness of the tone and noise [6]. In a recent paper presented at a conference of the European Society for the Cognitive Sciences of Music, Gregory Sandell distinguished between "emergent" timbre, produced by the perceptual fusion of the entire spectra of two separate instruments, and "augmented" timbre, produced by the fusion of several harmonics from one instrument with the entire spectrum of a second instrument, thus changing the timbre of the second while leaving the perceptual "distinctiveness" of the first intact [7]. Something of the same distinction might well be relevant here, where the flute timbre is augmented by the presence of noise, although the noise itself is not sufficiently intense to produce the kind of emergent timbre tested for in our experimentation. With the flute as an example, it appears that fusion ought to be defined from a perceptual point of view—i.e. in terms of the mutual influence of the noise and tone on the timbre of the other. Thus, we might say that noise fusion exists if two conditions are met: (1) the overall timbre of the sound must demonstrate the mutual influence of the two components (so that in the case of two of the instruments represented here, we have a noisy flute and a noisy sanza, as opposed to a flute and sanza in the presence of noise), and (2) when the sound is analytically divided into its two components, both the noise alone and the tone stripped of the noise must sound discernibly different in the absence of the other.

The second observation we can make from the flute spectra is that while the noise envelope does follow the harmonic envelope, beginning with the second harmonic, it is actually shifted slightly, such that noise peaks are higher in frequency than harmonic peaks. Furthermore, Fig. 9, which includes spectra of both the harmonics and the noise, shows what appears to be a second set of harmonics, beginning below the fundamental, whose peaks are in fact as regular as the real harmonics. Our noise-separation program assigns most of these peaks to the noise component. However, their regularity made us suspicious of the accuracy of this assignment until we matched the pitch of the tone and found it in fact to correspond to a frequency of about 1,030 Hz. These regular peaks, then, become the peaks of the noise, which are themselves quite regular harmonically. The shift of the noise center frequency is greater than that predicted by our experiments on the coincidence of tone frequency with the center frequency of the noise, which may contribute to the weaker influence of the noise on the flute timbre. At the same time, the *real* partials of this traditional flute are remarkably inharmonic for a flute [8], although not so much in comparison to the sanza discussed below. If the partials were ideally harmonic, their values would lie somewhere between their actual peaks and the harmonically regular peaks of the noise that surrounds them. Moore [9] and others have shown that especially within the first six harmonics of a complex tone, inharmonic partials tend to pull the sense of pitch away from that indicated by the frequency of the fundamental in the direction of the inharmonicity. If this is true, then a possible effect of the regular noise peaks located above the true harmonics might be the undoing of the tendency of the flute's (inharmonic) partials to pull the sense of pitch downward, by adding what might almost be considered a second set of partials, each paired with one of the real harmonics, but offset in a direction opposite to the real harmonics' inharmonicity. Indeed, in comparison with the complete tone (noise and harmonic frequencies) the harmonic-only signal (without noise) sounds lower. Even stronger confirmation of the role of noise in undoing the pitch effects of inharmonicity is heard in a comparison of the harmonic-only signal with the noise-and-harmonic signal when both have been filtered to exclude all but the first three harmonics. Informal listening to these sounds produced general agreement among listeners that not only was

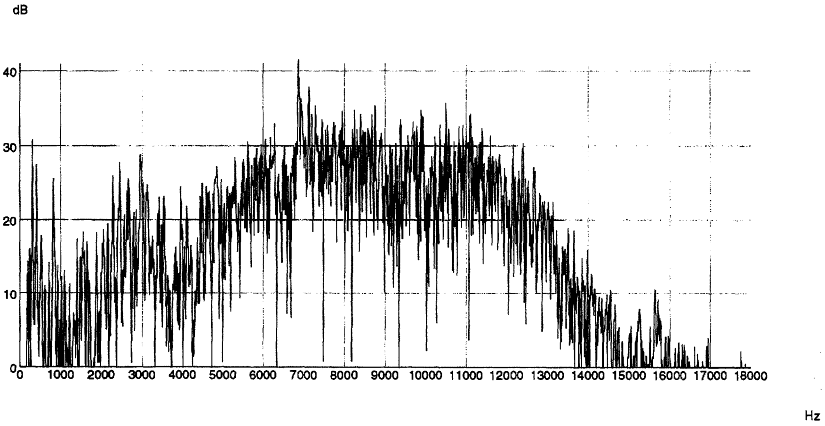


Fig. 13. Narrowband spectrum of musical bow tone (*complete*) of fundamental frequency 302 Hz. The horizontal axis corresponds to frequency (in Hz), and the vertical axis corresponds to amplitude (in dB). Notice the presence of two kinds of noise (see text): noise that surrounds primary harmonics, as in sanza and flute tones, and high frequency noise (above about 6,000 Hz) that appears to produce the layering audible in its timbre.

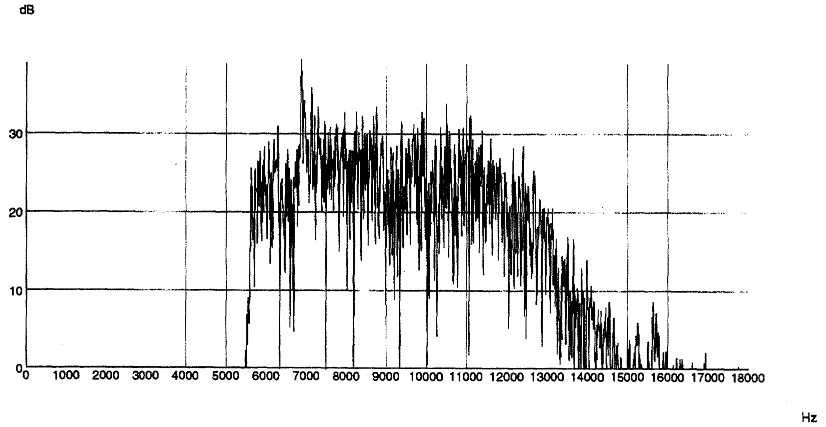


Fig. 14. Narrowband spectrum of high frequency noise filtered from spectrum. The horizontal axis corresponds to frequency (in Hz), and the vertical axis to amplitude (in dB). Notice the relative formlessness of the noise in comparison with the noise in the remaining signal (shown in Fig. 16).

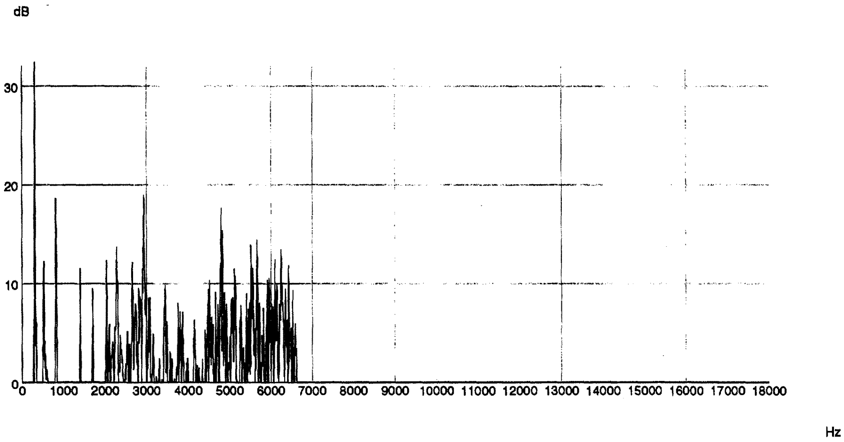


Fig. 15. Narrowband spectrum of portion of signal remaining (after high frequency noise was filtered from spectrum) with noise removed by noise-separation process. The horizontal axis corresponds to frequency (in Hz), and the vertical axis corresponds to amplitude (in dB).

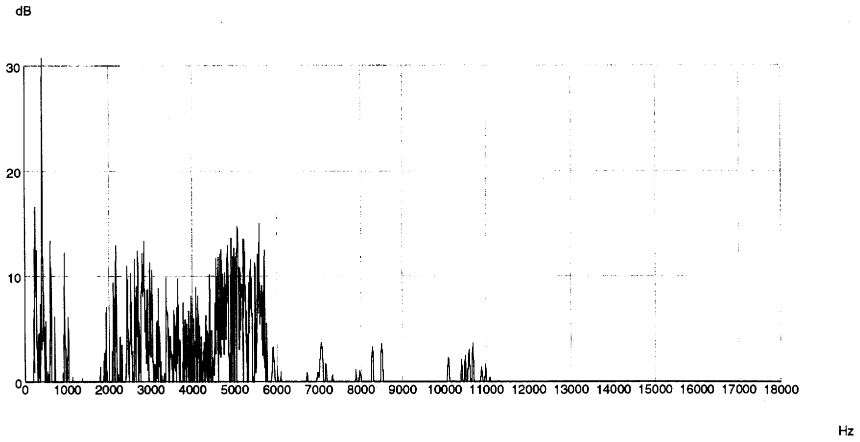


Fig. 16. Narrowband spectrum of portion of signal remaining (after high frequency noise was filtered from spectrum) with noise only, isolated by noise-separation process. The horizontal axis corresponds to frequency (in Hz), and the vertical axis corresponds to amplitude (in dB). Note, for this portion of the signal, the similarity in envelope (outline of peaks) of the noise spectrum to the harmonic spectrum above, indicating that noisebands are roughly centered on primary harmonics.

the pitch higher for the reduced noisy complex, but the "roughness" or "graininess" of the noiseless complex was reduced by the presence of noise.

Sanza Analyses

The spectra for the sanza are remarkably complex when compared with the flute spectra. The partials of the sanza have so inconsistent a relationship (the distance between them varying from 212 to 320 Hz, with a quasi-fundamental of 285 Hz) that it is difficult to decide which peaks are legitimate harmonics—even though the sample was run through the noise-separation process three times to eliminate as much noise as possible. Here again, the noise envelope follows the spectral envelope. Unlike the flute, however, the noise bandwidths surrounding the harmonics begin to approximate the size of the respective auditory filters only with the fifth harmonic. Further, with the exception of the fundamental, the intensity level of the noise relative to the harmonic it surrounds is high enough that the total noise power entering each filter appears to exceed the noise-to-harmonic ratio necessary for fusion.

A problem in the noise analysis of the sanza (which did not arise in the case of the flute) is that, in all likelihood, the instrument used for analysis here actually contains more than one noise source. We have mentioned already the use of bottle caps attached to the resonator of the instrument. At the same time, there is probably some degree of noise produced by the instrument itself, as it is traditionally performed. The noise-separation process, of course, cannot distinguish between the two sources. This is always a problem when there is a supplementary noise device acting on an instrument. For the analysis of overall noise content, it probably makes little difference where the noise originates, as long as it is discernibly fusing or layering. However, when discussing the reaction of supplementary noise devices on the "natural" timbre of the instrument, it is important to distinguish noise produced in the absence of the device from noise produced by the device. Thus, having suggested a possible "function" of the sanza noise, we also point out that propositions as to noise devices that "compensate" for inharmonicity, for example, must remain tentative as long as there is no sample of the same instrument's sound with its inherent noise and *without* the external noise device—that is, as long as there is no pure sample of what it is that the noise compensates for.

Musical Bow Analyses

As mentioned earlier, the musical bow example differs from the other instruments described here in that it demonstrates layered—rather than fused—tone and noise. When we tried our noise-separation process on samples of the bow, we found that the software left so much noise in the frequency region above 6,000 Hz that a spectrum of what was meant to be harmonic-only elements looked quite similar to the spectrum for the entire tone. When we next bandpass-filtered this extracted signal to leave only the region above 6,000 Hz, we were unable to hear any periodic elements whatsoever; it was evident from this result that the bow was an example of an instrument on which our noise-separation process is ineffective, since the noise elements in part of the spectrum appear to be louder than the harmonic elements in the same regions. Thus the software's choice of the highest intensity peaks as harmonics was inappropriate in the case of the musical bow. However, having filtered out the frequency region above 6,000 Hz, we could then apply the noise-separation process to the sample that remained. Figures 13–16 are spectra of the entire bow tone (Fig. 13), the high-intensity noise region removed by filtering (Fig. 14), the remaining signal with noise extracted (Fig. 15), and the noise only of the remaining signal (Fig. 16).

If we look first at the spectra of the signal remaining after filtering out all frequencies above 6,000 Hz, we see something of the same pattern of noise-tone relation as we saw in the flute and sanza. The noise follows the harmonic content of the tone, forming noise bands surrounding each formant peak. Again, when we further filtered this part of the signal to leave only the first four harmonics, those harmonics were inharmonic to the extent that they were heard to be so weakly fused that, rather than a single, unified sound with a discernible pitch, almost all listeners heard a collection of three or four pitches, none of which appeared to be the same pitch as the entire tone. When we combined these four harmonics with the *high frequency noise* (Fig. 13), the resulting percept was of the separate pitches of the harmonics with noise superimposed or layered above. Evidently the noise above 6,000 Hz contributed neither to the fusion of the harmonics nor to the sense of pitch but, rather, added an additional, layered sound. When we mixed back in the noise below 6,000 Hz (Fig. 17), however, listeners reported that the same four harmonics fused into a slightly noisy tone and yielded a sense of pitch matching that of the entire tone. Again, it appears that only the noise surrounding harmonics of the bow could aid in their fusion and the resulting pitch perception, and further, that this noise was also adding a slight noisy quality—comparable in intensity to the noisy quality of the flute. When we then mixed back in the high frequency noise (above 6,000 Hz), the resulting percept was almost identical to the entire tone as it occurs in performance, with a unified pitch and timbre-sense and an additional noise layered over this complex.

Thus, it appears that, in fact, the musical bow recorded for this sample produces two kinds of noise: one of lower frequency and intensity that fuses with the tones it surrounds and that contributes to their fusion and pitch sense, and another of higher frequency and intensity that produces the layering of noise audible above the tone. As in the case of the sanza, there is no way of knowing if both kinds of noise arise from the noise device attached to the instrument or if the fusing noise is inherent to the instrument, while the layered noise is the result of the "metal clackers." Whatever the sources of the noise, their configurations as revealed by the spectra above are again in agreement with the results of our experiments. The fusing noise centers roughly on prominent harmonics, while the layering noise is well set off from audible periodic elements and is of sufficient intensity that it masks any periodic elements that lie in the same frequency region.

One of the intriguing aspects of the study described in this paper—although it is one whose implications can only be speculated—occurred in informal responses from listeners describing the fused noise-tone percept. Four comments that were repeated frequently by subjects were that the noise-tone fusion (1) seemed to "spread out" the pitch, (2) gave the sensation of more than one pitch, (3) left listeners unsure of the pitch, and (4) caused the illusion of asynchrony between the noise and tone. Vincent Dehoux of the Department of Ethnomusicology in the laboratory of the Langues et Civilisations Traditionnelles Orales of the Centre National de Recherche Scientifique (LACITO-CNRS) in Paris tells the story of being in the presence of two sanza musicians from the Central African Republic who had been working for some time to tune their instruments together for a duet they were to perform. After a final, unsuccessful attempt, they gave up the effort, deciding instead simply to attach still more bottle caps to the resonators of their instruments. And, said Dehoux [10], whatever the mechanism, the bottle-cap trick appeared to work. It is in stories such as this one that the analytical and experimental parts of our noise fusion study come together. Anyone who is familiar with the noisy timbre of a bottle-cap–laden sanza such as the one described above will recognize that the level of the noise could not have been nearly sufficient to *mask* the mismatched intonation of the two sanzas. If there is any validity to our earlier suggestion that the noise described above counteracts the inharmonicity of the instruments that produce it—and if we consider the comments of our experimental subjects concerning the "spreading out of pitch" as an effect of noise—then it is possible that a similar phenomenon is working for the Central African sanza musicians to compensate for the imperfect intonation between the two instruments by blending one intonation into the other.

CONCLUSION

Verification of hypotheses concerning the perceptual effects of noise in regard to faulty tuning, inharmonicity or ambiguous pitch sensation will require further, extensive laboratory research. Verifying that such perceptual effects are intentional on the part of musicians will require additional extensive field research. Similarly, an investigation of the intended effect of noise fusion and layering—of whether one was more desirable or productive of the illusion than the other—will also require inventive field research. The use of other noise phenomena, such as auditory restoration effects, might also contribute to the intended effects of noise in African music. From the perspective of ethnomusicology, in order to combine the controlled results of laboratory research with the richness of uncontrolled musical behavior, experimentation of the kind described here must consist of equal parts of laboratory research, acoustic analyses of the instruments and field research.

Acknowledgments

This research was funded in part by National Science Foundation and Fyssen Foundation grants to the first author and by a grant from the French Ministry of Research and Space to the second author.

References and Notes

1. See, for example, J.H. Kwabena Nketia, *Music of Africa* (New York: W.W. Norton, 1973).

2. The sanza, also called *mbira* in parts of Africa, is a small, handheld lamellophone with a rectangular shallow box resonator and a clavier of wooden or metal lamellae that are "plucked" with the thumbnails. In some regions of Africa, the sanza may be put inside a larger calabasse for extra resonance. In the United States, the instrument is colloquially called a "thumb piano."

3. The musical bow is a struck string instrument resembling the bow of a bow-and-arrow with a gourd resonator attached to the wooden bow part of the instrument. This gourd is held against the musician's chest cavity for further resonance, while he or she strikes the string with a short baton.

4. For musicians, whose system begins counting harmonics after the fundamental frequency, this will be the first harmonic or overtone.

5. B.C.J. Moore, *An Introduction to the Psychology of Hearing* (London: Academic Press, 1989).

6. Degree of fusion was an issue of concern, in fact, in the original design of the experiment, but the intial difficulty encountered in proposing the notion of noise fusion to subjects made the proposition of *degrees* of fusion unrealistic.

7. Gregory Sandell, "Analysis of Concurrent Timbres with an Auditory Model," paper presented at the Third International Conference for Music Perception and Cognition (ICMPC), hosted by the European Society for the Cognitive Sciences of Music (ESCOM), Liege, Belgium, 23–27 July 1994.

8. Personal communication, Marc Pierre Verge (instrument acoustician, IRCAM) concerning inharmonicity of flute-like instruments, March 1994.

9. Moore [5].

10. Personal communication; see also Vincent Dehoux, *Les Chants à Penser* (Paris: SELAF, Place de la Sorbonne, 1992).

Bibliography

A.S. Bregman and S. Pinker, "Auditory Streaming and the Building of Timbre," *Canadian Journal of Psychology* **32** (1978) pp. 19–31.

B.R. Glasberg and B.C.J. Moore, "Derivation of Auditory Filter Shapes from Notched Noise Data," *Hearing Research* **47** (1978) pp. 103–138.

B.C.J. Moore, "The Perception of Inharmonic Complex Tones," in W. Yost and C. Watson, eds., *Auditory Processing of Complex Sounds* (Hillsdale, NJ: Lawrence Erlbaum Associates, 1987).

B.C.J. Moore, B. Glasberg and R. Peters, "Relative Dominance of Individual Partials in Determining the Pitch of Complex Tones," *Journal of the Acoustical Society of America* **77**, (1985) pp. 1853–1866.

B.C.J. Moore, R. Peters and R.W. Peters, "Thresholds for the Detection of Inharmonicity in Complex Tones," *Journal of the Acoustical Society of America* **77** (1985) pp. 1861–1867.