

Perception of musical tension for nontonal orchestral timbres and its relation to psychoacoustic roughness

DANIEL PRESSNITZER
IRCAM-CNRS, Paris, France

STEPHEN McADAMS
IRCAM-CNRS, Paris, France
and Université René Descartes, Paris, France

and

SUZANNE WINSBERG and JOSHUA FINEBERG
IRCAM-CNRS, Paris, France

Can tension in nontonal music be expressed without dynamic or rhythmic cues? Perceptual theories of tonal harmony predict that psychoacoustic roughness plays an important role in the perception of this tension. We chose a set of orchestrated chords from a nontonal piece and investigated listeners' judgments of musical tension and roughness. Paired comparisons yielded psychophysical scales of tension and roughness. Two experiments established distinct levels of these two attributes across chords. A model simulation reproduced the experimental roughness measures. The results indicate that nontonal tension could be perceived consistently on the basis of timbral differences and that it was correlated with roughness, the correlation being stronger as the perceptual salience of other attributes (such as high-pitched tones or tonal intervals) was reduced.

An impression of successive tension and release is a common experience while listening to music. The study of this intuitive notion is actually fundamental to Western music theory and has been a matter of debate for both music analysts and cognitive psychologists. Whereas many studies have helped us to understand tension perception in the tonal framework, the compositional practices of contemporary composers raise new questions. The use of a musical material that does not draw on the implicitly shared tonal culture may disorient listeners and, hence, deprive them of the essential landmarks that schemas of tension and release provide for the appreciation of a musical work. In this study, we propose to investigate the perception of tension levels for nontonal orchestral timbres by musician and nonmusician listeners. We hypothesized that if tension could actually be perceived in such a context, a basic psychoacoustic dimension known as roughness would account for an important part of that perception. Roughness is an auditory attribute that has

been proposed as a sensory basis for musical consonance within the tonal system (Helmholtz, 1877/1954).

Comparing tension and roughness for a set of stimuli assumes that there is a parallel between a complex musical concept and a basic sensory attribute. This requires a preliminary specification of the two terms and of the reasons for taking a new perspective in considering them jointly.

Musical Tension in a Tonal Context

The identification of *tension-release* schemas is the starting point for most classical harmonic analyses of Western music. The method proposed by the music theorist Heinrich Schenker (1935/1979), for instance, defines hierarchical structural levels that are, in the end, reducible to a fundamental *Ursatz*, which is very similar to a basic pattern of tension and release. The *generative theory of tonal music* (GTTM; Lerdahl & Jackendoff, 1983), which attempts to account for music perception and understanding by principles derived from cognitive psychology, holds the existence of a network of tension and release to be a fundamental axiom. The links it creates between events over time help us to hear a piece in a coherent way, rather than in the form of a juxtaposition of isolated sounds. These schemas have also been proposed as a basis for arousal and emotion, underlining another aspect of their importance in music listening by trained or untrained listeners (Langer, 1967; Meyer, 1956).

Part of this work was completed thanks to a Fyssen Foundation grant to the first author. The authors thank two anonymous reviewers for extremely useful comments on earlier versions of the manuscript. Correspondence concerning this article should be addressed to either D. Pressnitzer or S. McAdams, IRCAM, 1 place Igor Stravinsky, F-75004 Paris, France (e-mail: pressnitzer@ircam.fr or smc@ircam.fr).

—Accepted by previous editor, Myron L. Braunstein

There exists no simple, universally valid, and widely accepted definition of what tension and release actually are. To indicate the scope of the notion, let us take the basic example of the authentic cadence in Western tonal harmony. In this succession of two musical events, a dominant seventh chord based on the fifth degree of a major scale is considered unstable and dissonant, calling for a resolution to a tonic chord based on the first degree, which is stable and consonant. It is generally accepted that this instability constitutes a *tension*, its resolution a *release*. The combination of these two is the well-known pattern at the conclusion of the vast majority of tonal pieces written in the 18th and 19th centuries. However archetypal, this example embodies several aspects of the notion of musical tension.

Tension is thought to be, in part, linked to the intrinsic consonant or dissonant aspect of the chords. The dominant seventh chord contains a tritone interval, formerly known as *diabolus in musica* (considered to be dissonant, as one may guess), and induces tension. The tonic triad contains only a perfect fifth and a major third, two consonant intervals, and therefore induces very little tension. Attempts were made to ground the idea of intrinsic consonance in a feature of the acoustic world: the overtone series. That is to say, a set of frequency components with an integer multiple of a fundamental frequency has been proposed as a physical basis for the phenomenon (Mersenne, 1636/1975; Rameau, 1722/1971). Theories emphasizing perception then introduced the concept of roughness that we will discuss below (von Helmholtz, 1877/1954).

Tension can also be thought of in terms of being an expectation, the fulfillment of which brings a release. From this point of view, in the cadence example, the temporal succession of the two chords might play a role. The two notes of the tritone from the dominant chord are unstable and just one semitone apart from the tonic and the third of the tonic triad, which are stable. Therefore, the tritone is said to resolve on the major third. This idea of resolution, or good continuation, has been proposed as a musical parallel to the laws of Gestalt psychology (Meyer, 1956; Narmour, 1990).

Lastly, the arousal of the expectation could be linked to cultural habituation: The classic cadence pattern is recognized, and the resolution is expected. Neural nets have been able to reproduce the expectancies of listeners concerning tonal music after a preliminary training phase (Bharucha & Olney, 1989). The habituation consists in internalizing the organization of tonal pitch space (Krumhansl, 1990). In a study in which tension perception for both melodies and short chord sequences was investigated, Bigand (1993) confirmed the existence of explicit or implicit tonal acculturation.

Within the tonal framework, several factors may, therefore, have an influence on the perception of schemas of tension and release. These influences can be concurrent, as in the case of a perfect cadence, or contradictory, if the composer seeks some kind of structural ambiguity. In any

case, in the tonal framework, a set of established rules implicitly shared by the audience is available to express tension and release.

Musical Tension in a Nontonal Context

The problem is very different for a composer who does not wish to refer to tonality. Let us take, as an example, the genesis of the sequence chosen to be the basis for this study (score shown in Figure 1).

Spectral analyses were performed on recorded notes played on a musical instrument—in this case, a double-bass with rapidly changing vibrato depth, speed, and bow pressure. Sets of four to five partials, selected from among the strongest components of the analysis within a fixed band of frequencies, were isolated to create a reservoir of material. Eight chords from this reservoir form the sequence used in this study. The chords were arranged to create a tension pattern that was needed for musical reasons. Since the chords resulted from a raw acoustic analysis, no classic tonal function could be ascribed to them. Therefore, the composer chose to define an intrinsic tension value for each of them by means of a computational algorithm. The greatest common divisor of the constituent frequencies within a tolerance corresponding to a sixteenth of a tone was held to be a qualitative indication of tension: the lower the divisor, the more inharmonic and, therefore, tense the chord was considered to be. Figure 2 displays the output of the algorithm, each chord numbered along the abscissa having a computed value along the ordinate. The shape of the curve was also used compositionally to reinforce this tension contour with other musical parameters.

This approach raises several questions. First, is the notion of tension still relevant for material without any tonal function? Second, in the score, the movements are underlined by rhythmic and dynamic differences, whereas the tension algorithm used only average spectral domain data and did not take into account these supplementary cues. To be coherent, steady-state timbral differences between the individual chords should be able to convey the tension movements alone. The first aim of this study was, therefore, to establish whether differences in tension that were not based on tonal, rhythmic, or dynamic cues could actually be perceived by listeners for these orchestral timbres. In such an unfamiliar context, listeners would need to rely on universal psychoacoustic cues until style-specific cues could be learned.

Roughness

Such a psychoacoustic cue may be found in auditory *roughness*. Roughness is caused by the perception of rapid amplitude fluctuations in the range of 20–200 Hz. Simultaneous tones close in frequency can cause rough beats (e.g., a minor second played in the medium register of the piano). Experimental studies that have sought to quantify roughness perception have stressed its close relation to peripheral auditory mechanisms. Beating tone pairs for

The image displays a musical score for an orchestral excerpt from *Streamlines* (Fineberg, 1997). The score is organized into two main systems. The first system includes staves for Flute 1, Flute 2, Clarinette Solo, Violon, and Alto. The second system includes staves for Fl. 1, Fl. 2, Cl., Vln., and Al. The score is divided into sections labeled T1 through T8, which are chords recorded individually. Dynamic markings such as *mp*, *pp*, *f*, *mf*, and *p* are used throughout. Articulation symbols like accents and slurs are also present. The time signature changes from 3/4 to 4/4 and then to 3/4.

Figure 1. Score of the excerpt from *Streamlines* (Fineberg, 1997), studied in our experiments. The chords indicated at the top of the staff (T_1 , etc.) were recorded individually, with the same durations and dynamic patterns, and constitute the eight *orchestral timbres* used as stimuli.

which the frequency difference and mean frequency were systematically varied (Plomp & Levelt, 1965) caused maximum roughness to be perceived with a frequency difference approximately equivalent to 25% of the auditory critical bandwidth¹ (as measured by Zwicker, Flottorp, & Stevens, 1957). As a matter of fact, roughness had been

used to measure the critical band, since the activation patterns on the basilar membrane of two elements of a tone pair necessarily have to overlap in order to interact and produce roughness (see Greenwood, 1991, for a review). Further studies refined the quantitative results according to roughness calculations for complex tones and led to

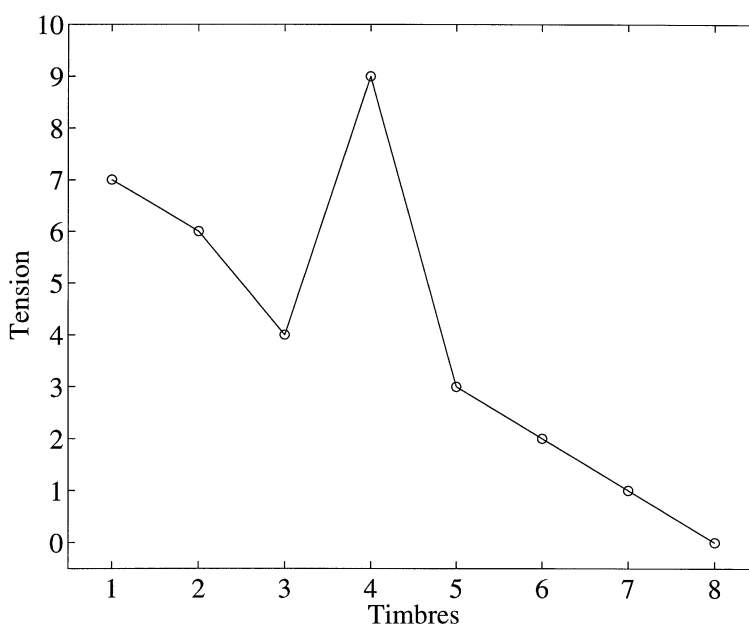


Figure 2. The tension measure originally computed by the composer. Timbres (T_1 , etc.) are shown on the abscissa in an order corresponding to their appearance in the musical score. The scale along the abscissa is not continuous, but the lines between the points have been added to reveal the tension profile over time. The tension scale is arbitrary and was used by the composer as a ranking.

computational models based on a sound's spectral composition (Hutchinson & Knopoff, 1978; Kameoka & Kuriyagawa, 1969a). The studies of amplitude-modulated tones (Terhardt, 1974), of amplitude-modulated noise (Fastl, 1977), or of frequency-modulated tones (Kemp, 1982) underlined the link of roughness perception with the low-pass characteristic of amplitude modulation coding by the auditory system (Viemeister, 1977). Other computational models have, therefore, been proposed that attempt to predict roughness for natural sounds in the temporal domain, using signal-processing techniques (Aures, 1985). Convergent approaches in the fields of developmental and comparative psychology have shown that young infants (Schellenberg & Trainor, 1996) or members of other species, such as birds (Hulse, Bernard, & Braaten, 1995), were more sensitive to differences in roughness of simple intervals than to differences in interval size. The finding that different species perceive roughness may be related to the fact that amplitude modulation coding does seem to be a basic feature shared by the auditory systems of both mammals and birds (Langner, 1992). Physiological correlates to roughness were also found in the discharge patterns of auditory nerve units (Tramo, 1996). These results confirm the existence of a basic sound attribute tightly linked to peripheral auditory mechanisms.

The musical repercussions of roughness have been addressed as well in some of these studies (Hutchinson & Knopoff, 1978; Plomp & Levelt, 1965; von Helmholtz, 1877/1954). The relative roughness of tonal intervals of

harmonic sounds was shown to correspond to their musical consonance. An interval of two simultaneous tones separated by a frequency ratio of 2:1 (one octave) produces no roughness, as all the partials of the upper note will correspond to partials of the lower note and will cause no beats. If the two tones are a just fifth apart (3:2 ratio), a few partials will be mistuned and will produce a little roughness. If the frequency difference is a tritone (45:32), a greater number of adjacent partials will cause beats and roughness. It follows that the relationship between frequency ratio and perceived consonance should be different if the simultaneous tones do not have harmonic spectra. This effect has been experimentally verified (Geary, 1980; Mathews & Pierce, 1980). Roughness has, therefore, been proposed as a sensory basis on which tonal harmony was built, allowing the development of more complex and style-specific rules.

However, most of the experimental results that support this proposition were obtained with simplified stimuli, and the experimental conditions (most often, direct judgments of *consonance* or *pleasantness*) used to obtain them could be questionable in a real musical context. The contribution of roughness to tension perception in a musical context has, nevertheless, been addressed in a study concerning short chord sequences (Bigand, Parncutt, & Lerdahl, 1996). Roughness, as estimated by model simulation, played a role in the tension judgments, together with other factors, such as implicit acculturation to the tonal function of the chords. In addition to study-

ing directly the perception of tension in nontonal harmonies, a second aim of this study was to gather roughness estimates of complex nontonal musical sounds and to investigate their contribution to tension perception.

EXPERIMENT 1

Method

Stimuli. A sequence of eight chords was chosen from *Streamlines* (Fineberg, 1997). The score of this sequence, simplified by the composer for the needs of the experiment, is shown in Figure 1. The chords were played in an ensemble setting (two flutes, clarinet, violin, and viola) by professional musicians who were members of the *Itinéraire* ensemble. These chords, sharing a homogeneous register, were recorded separately with the same duration and increas-

ing-then-decreasing dynamic pattern (*dal niente-mf-dal niente*). An example of the temporal and spectral characteristics of one of the stimuli is shown in Figure 3. For the sake of conciseness, the 14 other figures corresponding to the other stimuli are not reproduced here but would be very similar. Thus, the stimuli differed mainly in their timbral characteristics, and we will refer to them henceforth as *orchestral timbres*.

Apparatus. The recording took place in IRCAM's *Espace de Projection* concert hall. A pair of Schoeps microphones was placed 5 m from the instrumentalists at an elevation of 3 m. The stereo signals were transferred through a Neve V mixing table onto a Sony 7050 DAT recorder. They were subsequently digitally transferred and stored on a hard disk at a sampling rate of 44.1 kHz with 16-bit resolution through a Digidesign Protools interface.

The experiment used a Soluna S1 double-walled sound-insulated booth. The stimuli were played by a NeXT station through an ISPW

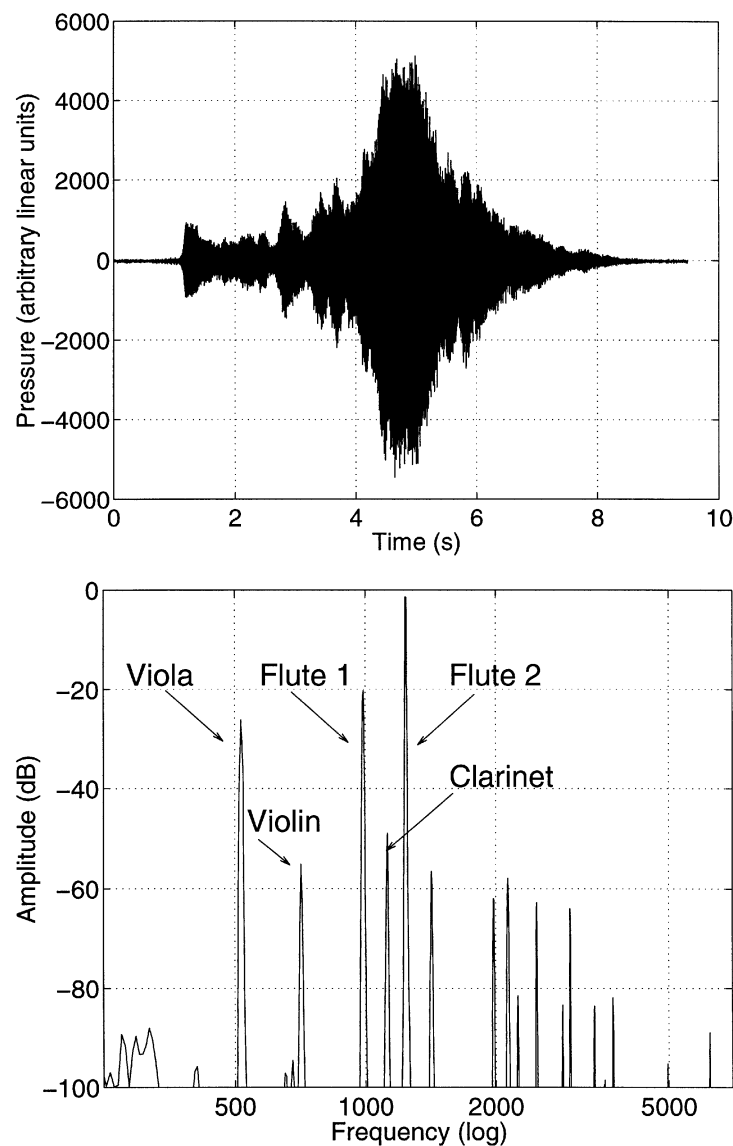


Figure 3. Waveform and long-term spectrum (obtained by Welch's method) of the orchestral timbre T_7 . The pressure unit, as well as the origin of the decibel scale, is arbitrary. The spectral components corresponding to the fundamental frequencies of the notes played by the different instruments are indicated.

DSP card, Pro IO digital-to-analog converters and power amplifier, and AKG K1000 open headphones. The measured sound level was approximately 80 dBA, adjustable at the request of the subjects so as to provide the most comfortable listening situation. The experiment was conducted in the PsiExp computer environment running on the NeXT workstation (Smith, 1995).

Subjects. The experimental group was composed of 30 listeners and did not include the authors. The subjects were 17–45 years old ($M = 26$). They came from widely different backgrounds, some working at IRCAM, others being recruited from a database of volunteers. The group included 9 professional musicians (composers or instrumentalists who had earned money making music over the last year), 8 amateur musicians (amateur instrumentalists), and 13 nonmusician listeners (with no musical training whatsoever). They were paid for their participation. One nonmusician listener did not finish the test and was subsequently removed from the results.

Procedure. The experiment included two parts, one concerning tension, the other concerning roughness. In the first part, the eight timbres were presented in all possible pairs in both orders without repetition (56 combinations). The subjects had to make a forced-choice judgment for each pair. The question they had to answer was, “Between the two sounds of this pair, do you hear an evolution from ‘tension to release’ or from ‘release to tension’” (translated from French *tension* and *détente*). They were allowed to define tension and release according to any criteria that seemed relevant. After

three to five test trials, in which no feedback was given, this first stage lasted approximately 30 min. In the second part, listeners’ attention was directed to one particular sound attribute, roughness. The concept of roughness was introduced through a typical sound example: a 1000-Hz sine wave that was amplitude modulated at 70 Hz, the modulation depth of which could be varied at will. This example was considered to produce no roughness if the modulation depth was zero, and the maximum roughness was obtainable with a single amplitude-modulated tone when the modulation depth was one (Zwicker & Fastl, 1990). The subjects were instructed that the difference in the quality of the tone they heard when moving the modulation depth slider was caused by roughness. Special attention was paid to not associating roughness with the *dissonant* and *uncomfortable* aspects of the stimuli in the instructions. The 56 pairs were then presented again in random order, and the subjects had to make forced-choice comparisons concerning changes in roughness. The two parts were always presented in this order (tension judgment, then roughness judgment) in order not to induce a bias in tension judgments because of prior knowledge concerning the notion of roughness.

Analysis methods. Using the forced-choice judgments as a basis, the eight timbres were ordered along a linear psychophysical scale, to allow a comparison with the composer’s predicted tension. A Bradley–Terry–Luce (BTL) analysis was performed (David, 1988). This method makes the assumption that each timbre has a *true* value

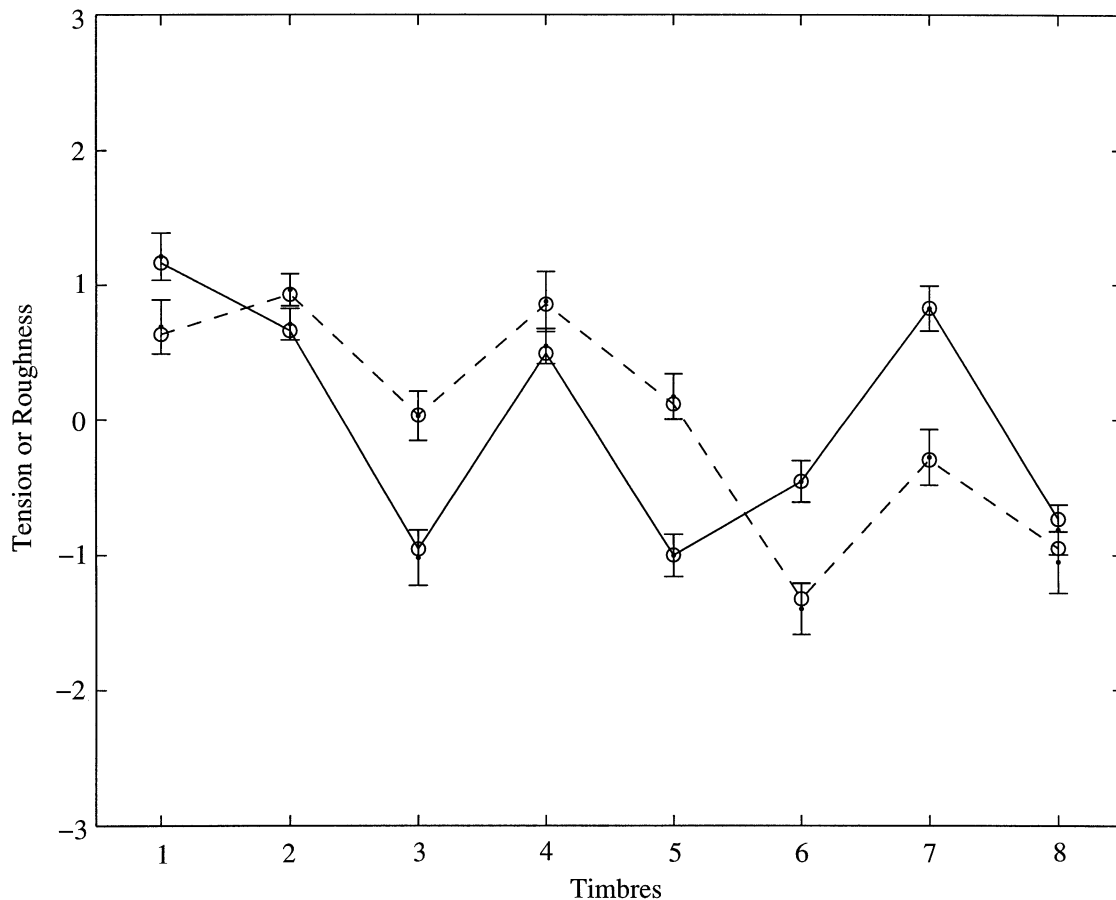


Figure 4. Psychophysical scales for tension (solid lines) and roughness (dashed lines) for Experiment 1. The timbres are labeled along the abscissa; the results of the BTL analysis and the standard deviations estimated by bootstrap are displayed along the ordinate. Empty circles represent the BTL values, and dots the means of bootstrap replications. The scales are not continuous, the lines between the points having been added for visual clarity. The scales are arbitrary and represent the contrast between items.

according to tension or roughness. It is this value that we try to estimate. During any experimental comparison between two timbres, listeners do not compare the true values but, rather, two associated random variables whose distributions are centered on the true values. In the BTL hypothesis, the distribution of the random variables is approximated by a hyperbolic secant law with constant variance (this yields a cumulative distribution similar to the cumulative normal, differing only at the tails). The proportions of judgments observed are then used to estimate the true values.

The result of the BTL analysis is a single tension or roughness value for each timbre. To test whether the subjects judged that stimuli indeed differed with respect to tension or roughness, the differences along the psychophysical scales were compared with the standard deviation (SD ; 68% of the central cases) of the results, as in an analysis of variance (ANOVA). However, an ANOVA assumes a normal error distribution, which is not necessarily the case here. A bootstrap analysis was therefore made, in which the SD s were estimated by using an empirical distribution obtained by means of resampling with replacement from the sample (Efron & Tibshirani, 1993). A bootstrap sample size of 100 was used, which is adequate for estimates of SD . The SD s are displayed in Figure 4 around the mean of the bootstrap values. The departure of the error from normality is noticeable by the asymmetry of the error bars about the BTL value.

A test of the null hypothesis between two items at a given risk α is usually achieved by examining the confidence intervals at $1 - 2\alpha$ attached to each of them. If the confidence intervals overlap, the null hypothesis cannot be rejected. Therefore, the statistical criterion we will adopt for inferring a significant difference between values at $p < .05$ is the comparison of the absolute difference between the values with the 90% confidence intervals, defined as $\pm 1.645 * SD$ (Efron & Tibshirani, 1993).

Results

The results are displayed in Figure 4. Let us consider first the tension measure (solid lines). The highest tension value was obtained by timbre T_1 , whereas the lowest value was produced by T_5 . On the BTL psychophysical scale, the absolute difference between these two extremes is more than an order of magnitude (10 times) bigger than the SD attached to the judgments. Using the significance criterion based on confidence intervals, the differences between the extremes of the psychophysical scale are significant. Therefore, all the stimuli were not randomly judged as being equally tense. Closer examination of the psychophysical scale for tension indicates that there were two broad groups: One, which included T_1, T_2, T_4, T_7 , was judged as being more tense, and the other, which included T_3, T_5, T_6, T_8 , as being less tense. The differences within the groups are not significant, whereas the difference between any pair of timbres across groups is significant.

The psychophysical scale for roughness (dashed lines) can be examined in a similar fashion. The difference between the most rough timbre, T_2 , and the least rough timbre, T_6 , is an order of magnitude greater than the SD s. Significant differences in roughness were, therefore, perceived between our complex orchestral sounds. The detailed examination shows that the roughness values were more regularly distributed along the psychophysical scale than were the tension values.

A comparison between the two psychophysical scales indicates that they differ for several timbres. With the non-

overlap of the 90% confidence intervals between the roughness and the tension values as the statistical inference criterion, differences are significant for T_3, T_5, T_6 , and T_7 . Specifically, T_6 and T_7 are judged as being tense but not rough, whereas T_3 and T_5 are judged as being rough but less tense.

Additional Analyses

Influence of presentation order. The order of presentation is not considered in the BTL analysis: Each pair of timbres was presented in both sequential orders to the listeners, on separate trials, and the judgments for the same pair were collapsed in a global proportion. The general influence of presentation order was, therefore, specifically tested to check whether it had an effect on listeners' judgments. An order sensitivity coefficient (osc) was defined as the percentage of trials in which the first timbre of a pair was chosen as being more tense or rough. If presentation order had no influence on judgments, the first element of a pair would be "chosen" as often as the second one across trials, and the coefficient should be equal to 50%. The statistical test for the osc was defined as follows: If the 50% value is included in the confidence interval at a given risk, order is considered as having no significant effect.

The osc is restricted to the interval $[0, 1]$, and the use of its SD to estimate a confidence interval would require a transformation. To avoid that, the confidence interval was estimated by bootstrap and 90% percentiles. A bootstrap sample size of 2,000 was chosen to estimate the confidence interval, defined as 90% of the results around the central case (Efron & Tibshirani, 1993). Presentation order was found to have a small but significant influence on tension judgments ($osc = .44$; $osc \in [.42, .47]$; $p < .05$). Order of presentation also slightly influenced roughness judgments ($osc = .47$; $osc \in [.46, .49]$; $p < .05$). Despite the marginal significance, the bias is small, and we interpret this analysis as an indication that order had only minor importance for both tension and roughness judgments. From a practical perspective, the bias is small enough to justify the BTL analysis that disregards presentation order.

Listening and/or judgment strategies. The examination of the agreement between listeners is a means to estimate whether all the subjects responded in the same way or whether different strategies appeared among them. This examination is all the more important because our subject group included professional musicians as well as amateur and nonmusician listeners, who were required to judge a complex notion such as tension. A relationship between strategies and subgroups of subjects having different musical education would indicate the influence of musical training on the experimental task.

A principal component analysis (PCA; Hotelling, 1933) was performed to study the tension judgments across subjects. Each listener was considered as a *variable* (as is the case in Q-factor analysis). The scores attributed to each timbre, defined as the sum of *more tense* judgments across the experimental trials for a given lis-

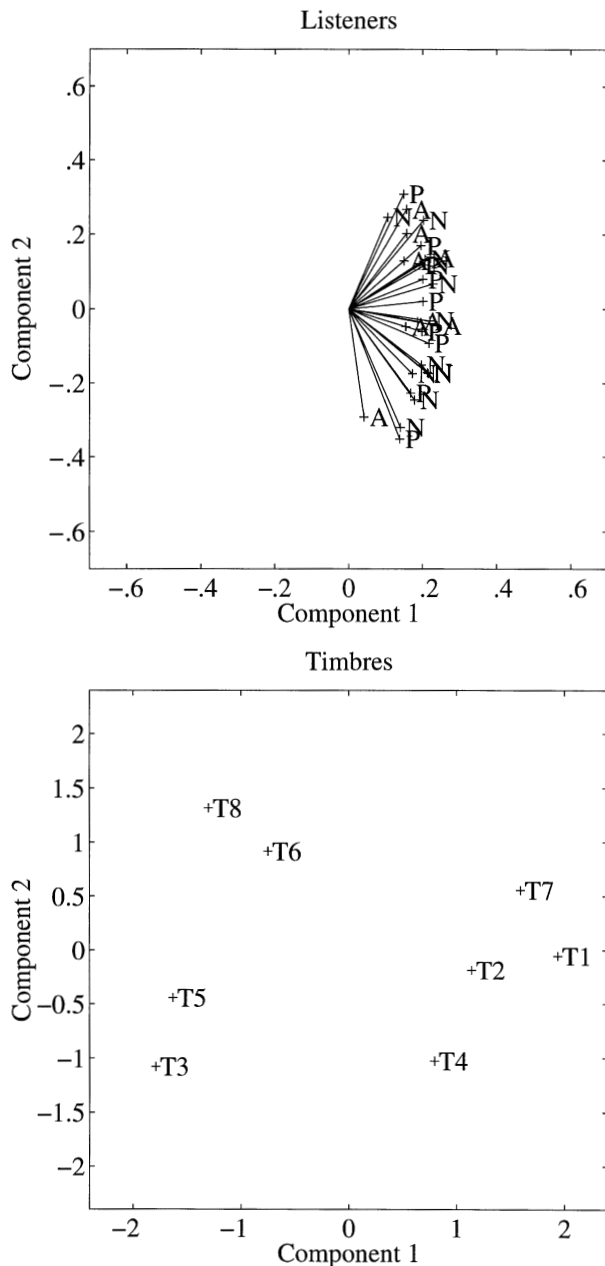


Figure 5. Results of the principal component analysis for the tension judgments of Experiment 1. Top panel: Listener coefficients for the two first principal components. Professional musicians are denoted as P, amateurs as A, and nonmusicians as N. Bottom panel: the component scores for each of the timbres are displayed. Timbres are denoted as T_1 to T_8 , as in Figure 4.

tener, were considered to be the *observations*. The correlation matrix between listeners could essentially be described by two principal components, accounting together for 75% of the variance. The other components, explaining less variance than did the original variables, are not considered any further.

The first component is, by definition, the main dimension of the data. It accounts here for 56% of the variance,

and all the subjects have comparable coefficients for this dimension (see Figure 5). In addition, the examination of the component scores shows that the order of the timbres from left to right corresponds to the ordering by increasing tension given by the BTL analysis (Figure 4). Thus, there is a consensus among the listeners as to the ordering of the timbres along the first dimension, since all the listeners have essentially the same coefficients with respect to this dimension.

The coefficients of the individual listeners are regularly spread along the second component, indicating differences among the subjects. However, the distribution of subjects according to their degree of musical training (P, professional; A, amateur; and N, nonmusician; see the procedure details) shows no particular cluster of one category. Therefore, this dimension cannot be explained in terms of differences in explicit musical education (as measured by our criterion). Concerning the timbre scores, the component opposes T_8 , T_6 , and T_7 to T_3 and T_4 . An interpretation for this component will be proposed in the General Discussion section.²

EXPERIMENT 2

Fusion and Specificities

Let us consider again the way the timbres were composed. Because of the use of spectral analyses of natural sounds, the score contains many notes that have quasi-harmonic relationships. For instance, for T_7 (see Figure 3), the viola and the flute are almost one octave apart. The harmonics of the different instruments overlap to a great extent. The feeling that derives from this technique, which is central to the composer's aesthetic, is that of *fusion* between the instruments. It would be difficult for us to define acoustically what is meant by the word *fusion*. Let us just point out qualitatively that a systematic interrogation of the subjects after Experiment 1 showed that none of them was able to name the five instruments used. Lack of segregation has been shown to be related, for example, to difficulty in identifying simultaneously presented vowels (de Cheveigné, McAdams, Laroche, & Rosenberg, 1995).

However, careful listening to the timbres reveals the presence of salient features in some of them. Timbres T_6 , T_7 , and T_8 contain a high flute note that was to be played at a moderate dynamic (see the score in Figure 1). In the recorded performance, these notes emerge from the timbres (see, for instance, the amplitude of the Flute 2 component in Figure 3) and give rise to what we call a *surface specificity*. For T_3 and T_4 , intervals with a harmonic function can be discovered in the score (for instance, in T_3 , a perfect fifth between the flutes and an almost major third between the clarinet and the second flute). These intervals evoke a strong syntactic function in the tonal system. In the composer's original intention, they should have been blended into the global timbre, but the relative balance between the instruments made them salient in our recordings. This constitutes what we call a *harmonic specificity*.

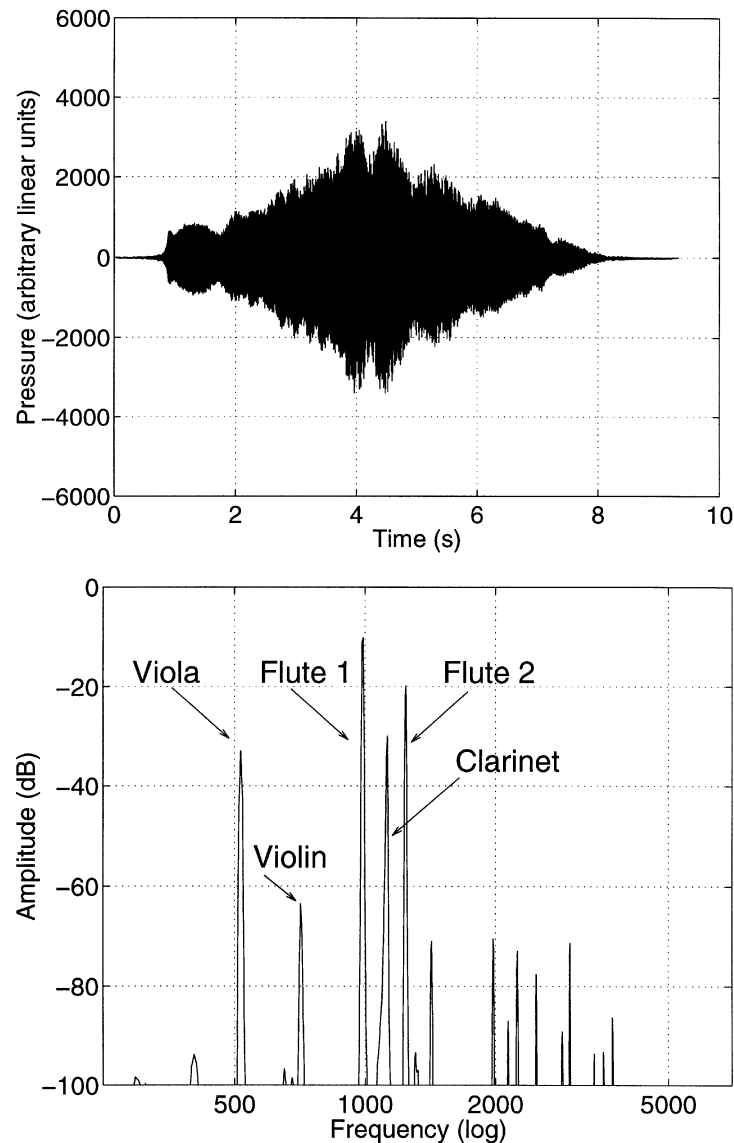


Figure 6. Waveform and long-term spectrum of the remixed timbre T_7 . The pressure unit, as well as the origin of the decibel scale, is arbitrary. The spectral components corresponding to the fundamental frequencies of the notes played by the different instruments are indicated.

A new experiment was designed in which modifications of the acoustic stimuli were done under the composer's supervision to obtain better fusion, according to his vocabulary. A multitrack recording was used to make a new version of the timbres in which the relative level of each instrument, as well as the evolution of the levels over time, was modified. Again, no systematic acoustic criterion for fusion can be proposed. Nevertheless, the acoustic analyses of the new version of T_7 are shown as an example in Figure 6. In this case, the remixing had the effect of correcting the balance between the instruments, most notably between Flute 1, Clarinet, and Flute 2. The waveform also displays a more regular envelope, the attacks of the different instruments having been smoothed.

This removes a potential cue for auditory segregation (Bregman, 1990). After the remixing of the eight timbres, a procedure similar to that in Experiment 1 was used to establish psychophysical scales for tension and roughness.

Method

Stimuli. New versions of the timbres employed in Experiment 1 were used. The same performances were selected, but the relative balance between the instruments' volumes, as well as their temporal evolution, was modified by remixing. An artificial room effect was added with a Lexicon processor, to reproduce the natural room color present in the stimuli of Experiment 1.

Apparatus. The multitrack recording on a 33.24A Sony multitrack recorder was made with five Neuman KM140 cardioid microphones, placed 50 cm from each performer's instrument. The

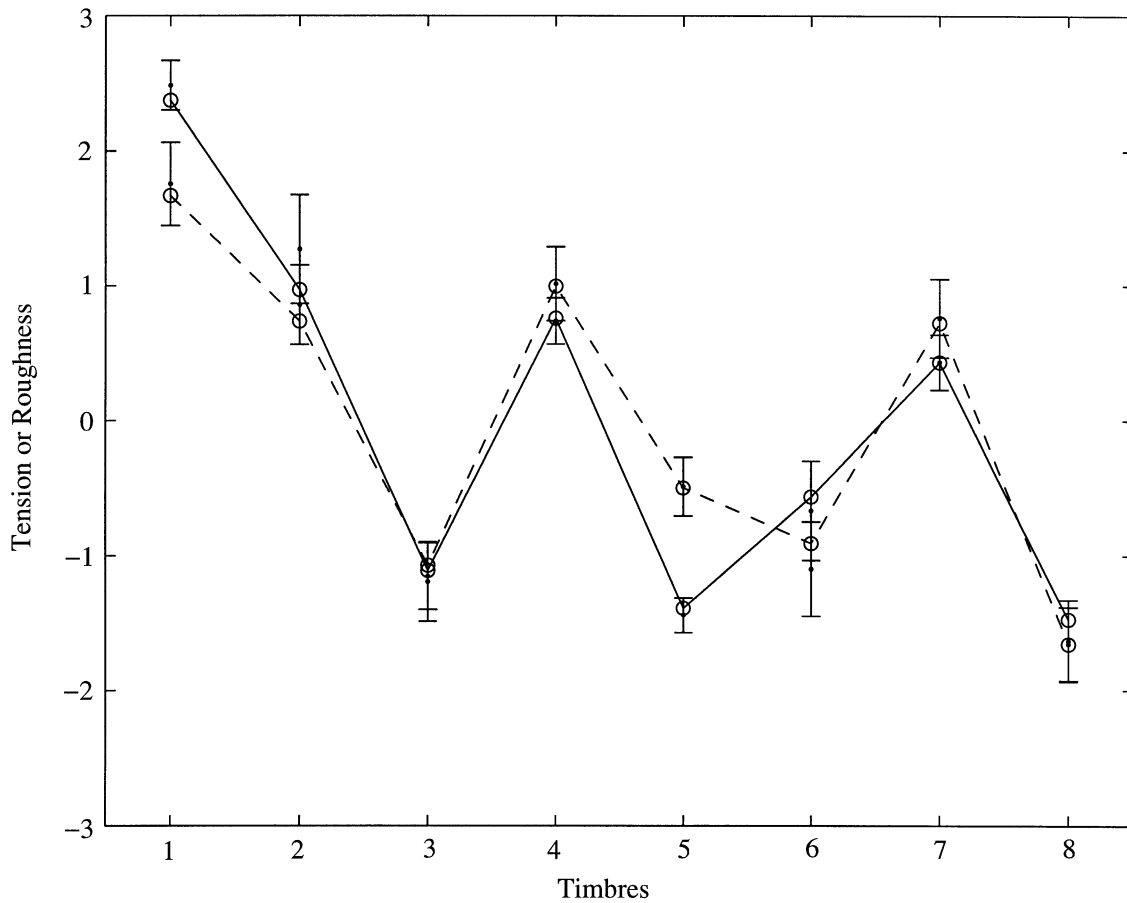


Figure 7. Psychophysical scales for tension (solid lines) and roughness (dashed lines) for Experiment 2. The timbres are labeled along the abscissa; the results of the BTL analysis and the standard deviations estimated by bootstrap are displayed along the ordinate. Empty circles represent the BTL values, and dots the means of bootstrap replications. The scales are not continuous, the lines between the points having been added for visual clarity. The scales are arbitrary and represent the contrast between items.

remixing was done through a Neve V mixing table to a Sony DAT 7050 player. The apparatus to play back the stimuli and to collect results remained identical to the one used in Experiment 1.

Subjects. A group of 10 subjects, from 18 to 45 years of age ($M = 27$), participated in the experiment. The subjects had not participated in Experiment 1 and were recruited in the same manner. The experimental group included 2 professional musicians, 4 amateur musicians, and 4 nonmusician listeners. They were paid for their participation.

Procedure. The experimental procedure remained strictly identical between the two experiments.

Results

The methods of analysis defined in Experiment 1 were used, and the results are shown in Figure 7. For tension judgments, the differences of the stimuli giving extreme values (now T_1 and T_8) are more accentuated than in Experiment 1, and the difference (according to the criterion defined in Experiment 1) is still significant ($p < .05$). The general response pattern remains similar: A distribution into two broad groups is obtained, with the exception that T_1 is now significantly more tense than all the

others. Therefore, the subjects could now establish more than two tension levels. Within the groups, the rank ordering of T_7 and T_8 decreased, whereas that of T_3 increased. However, the psychophysical scale for tension still differs from the one predicted by the composer's algorithm, especially for timbres T_6 , T_7 , and T_8 .

On the psychophysical scale for roughness, T_1 becomes the most rough timbre, and T_8 the least rough. The difference in their BTL values is larger than in Experiment 1, although the *SDs* are also larger. However, the difference is still significant. The general pattern remains similar, but the rank ordering of T_6 and T_7 increased.

The consequence of these changes in both psychophysical scales is that they now overlap to a much greater extent, in comparison with the results obtained from Experiment 1. Only one timbre, T_5 , displays a significant difference between tension and roughness values. A correlation coefficient was computed to describe the resemblance between tension and roughness judgments. The confidence interval of the correlation coefficient (restricted to $[-1, 1]$) was estimated by bootstrap and 90%

percentiles. The result in this second experiment was $r = .94$; $r \in [.84 .96]$, whereas the first experiment gave $r = .53$; $r \in [.32 .73]$. With the exception of the significant difference for T_5 (for which we find no plausible interpretation), these results suggest that roughness could now be sufficient to explain tension perception of the remixed timbres.

Replication of the Additional Analyses

The additional analyses performed on the results of Experiment 1 were repeated for the new experimental data. Order of presentation was found to have significant influence neither on tension judgments nor on roughness judgments.

The PCA analysis revealed a quasi-unidimensional solution, with a first consensus component that reflected the BTL psychophysical scale explaining 73% of the variance. The second component, accounting for only 15% of the variance, explains not much more variance than do any of the original variables. This indicates that, in this second experiment, the tension judgments could almost exclusively be explained in terms of the consensus dimension.

MODEL SIMULATION

A computational model of roughness (Hutchinson & Knopoff, 1978) was tested to try to reproduce the experimental results. The model considers a set of frequencies and amplitudes (f_i, a_i) intended to represent the spectral composition of a given sound. For each possible pair (i, j) of components, a partial roughness r_{ij} is computed according to their frequency difference. All the partial roughnesses, normalized by the amplitudes, are then summed

to give the global roughness that constitutes the model output. The corresponding analytic formula developed by Parncutt can be found in Bigand et al. (1996).

This model has theoretical weaknesses. The calculation of roughness between two frequency components is actually derived from experimental results (Plomp & Levelt, 1965), but the influence of differences in amplitudes between interacting components, as well as the simple addition between partial roughnesses, is not. This extrapolation of data concerning equal amplitude tone pairs to tone complexes is likely to be incorrect (Kameoka & Kuriyagawa, 1969a; Vos, 1986). This model was nevertheless tested, because its input is in a form similar to that of the composer's original algorithm. More generally, it is well adapted to musical notation, where chords can be represented by a list of component notes (frequencies).

To simulate our empirical findings, the data fed into the model should be the notes written in the score, with all the amplitudes equal. However, a spectrum also has to be associated with each note, because beats between partials are an essential source of roughness (Kameoka & Kuriyagawa, 1969b). Therefore, hypothetical spectra with four partials and a 12 dB/oct rolloff were chosen. This represents a schematization of the stimuli but remains plausible, considering the register and instrumentation used (see Figures 3 and 6).

The computed roughness values are shown in Figure 8. The correlation between the BTL analysis of experimental roughness judgments of Experiment 1 and computed roughness is ($r = .83$; $r \in [.73 .90]$, 90% confidence interval estimated by bootstrap percentiles). For the roughness psychophysical scale obtained in Experiment 2, the correlation is ($r = .88$; $r \in [.81 .94]$). The ex-

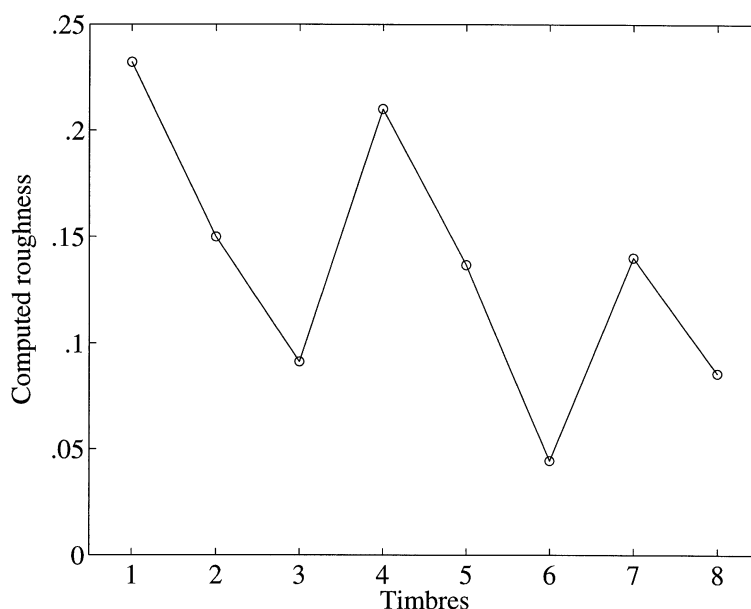


Figure 8. Model calculation of roughness for the eight timbres. The roughness scale is arbitrary.

perimental roughness judgments are, therefore, generally concordant with the model simulation.

With respect to using the model as a compositional tool, it is also possible to qualitatively compare its output with the composer's original tension predictions, shown in Figure 2. Considering roughness as a tension indicator, the model is globally better at predicting the experimental judgments; timbre T_7 , for instance, was predicted as bearing a low tension value but was experimentally judged as being both tense and rough. The model replicates this finding.

GENERAL DISCUSSION

Tension and Roughness

The first experiment allowed us to exhibit significant differences in tension for nontonal orchestral timbres. The notion of tension—deliberately defined loosely in the experimental instructions—gave rise to a consensus among subjects in a musical context that was unfamiliar to most of them. Roughness was also found to be a relevant criterion by which to judge the timbres. However, significant differences appeared between the two measures, indicating that factors other than roughness participated in tension judgments. The second experiment confirmed the perception of tension and roughness differences between the timbres but exhibited almost identical psychophysical scales. The correlation observed indicates that, in this second experiment, it is possible to account for the tension judgments by comparisons of roughness, increasing roughness inducing greater tension. Of course, a correlation is by no means a proof of causality, but other elements that appeared in the additional analyses are coherent with the hypothesis that roughness was a sensory basis for the perceived tension of the timbres.

The temporal succession of the timbres did not have a large influence on the tension judgments. Therefore, it is mainly the intrinsic aspect of the stimuli, which parallels the tonal concepts of intrinsic consonance or dissonance, that influenced the tension judgments. Roughness is thought to participate in intrinsic consonance in the tonal context, and our results extend this interpretation to a nontonal context.

The influence of the explicit or the implicit acculturation of listeners (all having a “Western” background) did not appear to be influential in a straightforward manner. That is, no correlation between musical education and judgment strategies could be found. Roughness, as a basic feature of sound, is not expected to depend on cultural factors and is readily available to trained or untrained listeners. However, this might also be due partly to the fact that musicians trained almost exclusively in tonal music may not have a special strategy for nontonal stimuli.

Finally, the correlation between psychophysical scales for tension and roughness is not due to a possible confusion between the words *tension* and *roughness*. In the experimental instructions, the definition of *tension* was left to the appreciation of the listeners, but *roughness* was defined by means of an example with sound. Roughness

estimations were then obtained in two distinct ways (an experimental evaluation and a model simulation) that gave converging results. Lastly, the significant differences observed in Experiment 1 between tension and roughness show that the subjects were indeed able to differentiate tension and roughness.

Comparison Between the Two Experiments

Experiment 1 displayed significant differences between psychophysical scales for tension and roughness, differences that generally disappeared in Experiment 2. Our interpretation of this is based on the specificities mentioned earlier. Although the status of these specificities is hypothetical, this hypothesis allows us to account for several aspects of the data.

Except for T_5 , all the timbres that displayed significant differences for tension and roughness in Experiment 1 contained specificities. Moreover, the specificities are coherent with the results: T_6 and T_7 were tense but less rough, and they included a prominent high-pitched note that could induce tension without roughness; T_3 contained intervals associated with release in the tonal system and was judged to be rough but less tense. These features, in addition to roughness, may have influenced the tension judgments. The effect of the remixing can then be interpreted as having reduced their perceptual salience. As fewer specificities emerged from the timbres, roughness became the main available cue to judge tension. The PCA performed to analyze listening and/or judgment strategies in Experiment 1 exhibited a secondary component contrasting subjects; this component opposes timbres containing a *surface* specificity to the ones containing a *harmonic* specificity. This component can, therefore, be interpreted as a disagreement in listening and/or judgment strategies concerning the impact of the specificities on tension perception. After the remixing of the timbres in Experiment 2, the second component disappeared.

Let us, nevertheless, point out that Experiment 2, in our view, is by no means an improved version of Experiment 1 in which possible “flaws” in the stimuli were discovered and then fixed. Our point is not to assert that tension and roughness are equivalent. In the general case, as was shown by Experiment 1, the influence of roughness combines with several other factors to contribute to the feelings of tension and release experienced by the listeners. The transformations made to obtain Experiment 2 stimuli, which basically consisted in homogenizing the timbres, suggest that roughness may *potentially* express tension in a situation that remains musically realistic. The comparison between the two experiments allows us to specify the conditions under which roughness can become a dominant factor in tension perception.

Remarks on the Experimental Procedure and Analysis Methods

This study has tried to draw a link between a basic psychoacoustic feature, roughness, and the musically important notion of tension. The choice of the stimuli, therefore, reflects a tradeoff between musical realism

and limitation of the number of experimental variables. On the one hand, the stimuli were extracted from an existing musical piece. They were recorded with live instruments in a concert-like situation and, in a second condition, were remixed in a situation habitually found in the preparation of recorded music. The question of finding a tension hierarchy was motivated by a preoccupation entertained by the composer during the writing of the piece. On the other hand, the timbres were homogeneous in register and dynamics, which may represent a limitation with respect to other musical situations. For instance, one may wonder how salient the tension cues linked to roughness would still be when placed in competition with cues of a different nature, such as rhythm, dynamics, or registral extremes. The interactions of melodic and harmonic factors were also not addressed here. These interactions are important in the context of polyphonic music, where it has been claimed that auditory streaming processes influence the perception of roughness and tension (Wright & Bregman, 1987).

Forced-choice paired comparisons were used to avoid direct magnitude estimation judgments. These judgments would have been hard to define without introducing terms not necessarily relevant to a musical context. Furthermore, the complexity of the stimuli could have made the task difficult for listeners. This design led us to adopt nonparametric statistical analysis methods in order to make as few hypotheses on the outcome of the data as possible: The bootstrap technique was used to examine the stability of the BTL results.

The analyses attributed a single tension value to each timbre. Of course, this is not supposed to mean that there exists an *absolute* intrinsic tension for the timbres investigated. It should be emphasized that the hierarchy is relative. The individual estimated tension values can be thought of in a fashion similar to that of the relative consonance and dissonance values traditionally ascribed to individual tonal chords. Although the level of “acceptable” dissonance has varied over the history of tonal music, it does not preclude perception of sensory consonance as a potential basis for tension perception (Bregman, 1991). Schoenberg (1911/1978), for example, preferred to refer to dissonances as “remote consonances”!

Perception of Tension in Nontonal Music

Theories of nontonal tension are far less numerous than theories of tonal tension. On the one hand, composers have proposed theories of nontonal tension that often have a close relationship to their own musical styles, as a codification of personal practices. A theory that is more generally relevant and, therefore, concerned with experimental testing is the extension of the GTTM to atonal music by Lerdahl (1989).

The GTTM reduces a musical piece—at different levels—to structural units by way of a *time-span reduction*. The reduction is made according to stability rules that are based on the hierarchical relationships that exist in the tonal pitch space. This reduction is then supposed to be the basis of the perception of tension and release move-

ments (*prolongational reduction*). For atonal music, Lerdahl (1989) makes the assumption that the pitch space is flat and replaces the stability rules by salience rules to obtain similar reductions. Both the tonal and the atonal time-span reduction theories have been experimentally tested, with contrasting results (Dibben, 1994). Whereas listeners were able to match a reduction to its original for tonal pieces, results were close to random in the atonal condition. The author proposed several reasons for this latter outcome, but the suggested conclusion is that the inability to match original pieces and reduced versions is inherent to the music tested (excerpts from Schoenberg opuses 11 and 19). Hierarchical levels of structure would be impossible to extract when listening to atonal music, and the only tension and release movements perceptible in this context (thanks to associations between sound events) would, therefore, be of lesser importance (Dibben, 1996).

Our results bring new elements to bear on this conclusion. The experimental judgments of orchestral timbres showed that tension and release movements could actually be perceived by listeners in a musical context not related to tonality. Another study gave a converging set of results by exhibiting a similar range of variation for tension judgments within tonal and nontonal musical excerpts and a significant effect of timbre on the judgments (McAdams & Pressnitzer, 1996; Paraskeva & McAdams, 1997). A difference between the present study and those of Dibben (1994, 1996) is that, here, the underlying structural harmony was approached in terms of timbre, rather than in terms of notated pitch events.³ Timbral quality alone, without a reference to a pitch space organization, was sufficient to convey intrinsic movements of tension and release between pairs of sounds. The question is now whether the local properties investigated (such as roughness) are able to convey tension and release movements on larger time scales. It would be of great interest to test whether originals and reductions, similar to the ones proposed by Lerdahl (1989) but based on timbral salience rules, including roughness, could be matched by listeners.

Timbre as potential form-bearing structure has interested composers, from the experiments of Berlioz and Debussy to Schoenberg's *Klangfarbenmelodie*, and more recently, those who belong to the *musique spectrale* trend (Murail, 1984; Pressnitzer & McAdams, 1999). Cognitive psychology has expressed the condition upon which a multidimensional attribute such as timbre could be the basis for building musical form—mainly in terms of the possibility of categorization and recognition of vectors in a timbre space (McAdams, 1989; McAdams & Cunebile, 1992). Our study is an indication that timbre may also be a privileged support for harmony, thanks to the dimension of roughness.

CONCLUSION

Two experiments have shown that the fundamental musical concept of *tension-release* schemas can be con-

veyed by orchestral timbres, even when no harmonic function corresponding to the Western tonal system can be defined for them. Different factors influenced these schemas: tonal references, heard both by musicians and nonmusicians, or surface characteristics resulting from performance variables, such as the emergence of high-pitched tones. However, when these features were reduced in salience, roughness became a noteworthy dimension. This sound attribute, in close relationship with peripheral auditory mechanisms, was highly correlated with the musical notion of tension. Its simulation with a predictive model may be considered. Roughness, as a dimension of timbre, could open up a way of thinking of non-tonal harmonies on the basis of perceptual phenomena.

REFERENCES

- AURES, W. (1985). Ein Berechnungsverfahren der Rauigkeit [A roughness calculation method]. *Acustica*, **58**, 268-281.
- BHARUCHA, J. J., & OLNEY, K. L. (1989). Tonal cognition, artificial intelligence and neural nets. *Contemporary Music Review*, **4**, 341-356.
- BIGAND, E. (1993). The influence of implicit harmony, rhythm and musical training on the abstraction of tension-relaxation schemas in tonal music phrases. *Contemporary Music Review*, **9**, 123-137.
- BIGAND, E., PARNCUTT, R., & LERDAHL, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception & Psychophysics*, **58**, 125-141.
- BREGMAN, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- BREGMAN, A. S. (1991). Timbre, orchestration, dissonance, et organisation auditive. In J. B. Barrière (Ed.), *Le timbre, métaphore pour la composition* (pp. 204-215). Paris: Christian Bourgois.
- DAVID, H. A. (1988). *The method of paired comparisons* (2nd ed.). London: Charles Griffin & Company.
- DE CHEVEIGNÉ, A., MCADAMS, S., LAROCHE, J., & ROSENBERG, M. (1995). Identification of concurrent harmonic and inharmonic vowels: A test of the theory of harmonic cancellation and enhancement. *Journal of the Acoustical Society of America*, **97**, 3736-3748.
- DIBBEN, N. (1994). The cognitive reality of hierarchical structures in tonal and atonal music. *Music Perception*, **12**, 1-25.
- DIBBEN, N. (1996). Perceptual facilities in atonal music: Implications for the listening experience. In B. Pennycook & E. Costa-Gioni (Eds.), *Proceedings of the Fourth International Conference on Music Perception and Cognition* (pp. 55-58). Montreal: McGill University.
- EFRON, B., & TIBSHIRANI, R. J. (1993). *An introduction to the bootstrap*. London: Chapman & Hall.
- FASTL, H. (1977). Roughness and temporal masking patterns of sinusoidally amplitude modulated broadband noise. In E. F. Evans & J. P. Wilson (Eds.), *Psychophysics and physiology of hearing* (pp. 403-415). London: Academic Press.
- FINEBERG, J. (1997). *Streamlines*. Paris: Éditions Max Eschig.
- GEARY, J. M. (1980). Consonance and dissonance of pairs of inharmonic sounds. *Journal of the Acoustical Society of America*, **67**, 1785-1789.
- GREENWOOD, D. D. (1991). Critical bandwidth and consonance in relation to cochlear frequency-position coordinates. *Hearing Research*, **54**, 164-208.
- HELMHOLTZ, H. L. F. VON (1954). *On the sensations of tone as the physiological basis for the theory of music* (2nd ed.; A. J. Ellis, Trans.). New York: Dover. (Original work published 1877; trans. from 4th ed., 1885)
- HOTELLING, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, **24**, 417-441, 498-520.
- HULSE, S. H., BERNARD, D. J., & BRAATEN, R. F. (1995). Auditory discrimination of chord-based spectral structures by European starlings (*Sturnus vulgaris*). *Journal of Experimental Psychology: General*, **124**, 409-423.
- HUTCHINSON, W., & KNOPOFF, L. (1978). The acoustic component of Western consonance. *Interface*, **7**, 1-29.
- KAISER, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**, 187-200.
- KAMEOKA, A., & KURIYAGAWA, M. (1969a). Consonance theory: Part I. *Journal of the Acoustical Society of America*, **45**, 1451-1458.
- KAMEOKA, A., & KURIYAGAWA, M. (1969b). Consonance theory: Pt. II. *Journal of the Acoustical Society of America*, **45**, 1459-1469.
- KEMP, S. (1982). Roughness of frequency-modulated tones. *Acustica*, **50**, 126-133.
- KRUMHANSL, C. L. (1990). *Cognitive foundations of musical pitch*. Oxford: Oxford University Press.
- LANGER, S. K. (1967). *Mind: An essay on human feeling* (Vol. 1). Baltimore: Johns Hopkins University Press.
- LANGNER, G. (1992). Periodicity coding in the auditory system. *Hearing Research*, **60**, 115-142.
- LERDAHL, F. (1989). Atonal prolongational structure. *Contemporary Music Review*, **4**, 65-87.
- LERDAHL, F., & JACKENDOFF, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
- MATHEWS, M. V., & PIERCE, J. R. (1980). Harmony and nonharmonic partials. *Journal of the Acoustical Society of America*, **68**, 1252-1257.
- MCADAMS, S. (1989). Psychological constraints on form bearing dimensions in music. *Contemporary Music Review*, **4**, 181-198.
- MCADAMS, S., & CUNIBILE, J. (1992). Perception of timbral analogies. *Philosophical Transactions of the Royal Society of London: Series B*, **336**, 383-389.
- MCADAMS, S., & PRESSNITZER, D. (1996). Psychoacoustic factors to musical tension in Western nontonal music [Abstract]. *International Journal of Psychology*, **3**, 148.
- MERSENNE, M. (1975). *Harmonie universelle*. Paris: S. Cramoisy. (Original work published 1636)
- MEYER, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.
- MURAIL, T. (1984). Spectra and pixies. *Contemporary Music Review*, **1**, 157-171.
- NARMOUR, E. (1990). *The analysis and cognition of basic melodic structures: The implication realisation model*. Chicago: University of Chicago Press.
- PARASKEVA, S., & MCADAMS, S. (1997). Influence of timbre, presence/absence of tonal hierarchy and musical training on the perception of musical tension and relaxation schemas. In *Proceedings of the 1997 International Computer Music Conference* (pp. 438-441). Thessaloniki: Aristotle University of Thessaloniki.
- PLOMP, R., & LEVELT, W. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, **38**, 548-560.
- PRESSNITZER, D., & MCADAMS, S. (1999). Acoustics, psychoacoustics and spectral music. *Contemporary Music Review*, **19**, 33-60.
- RAMEAU, J. P. (1971). *Treatise on harmony reduced to its natural principles* (P. Gosset, Trans.). New York: Dover. (Original work published 1722)
- SCHARF, B. (1970). Critical bands. In J. V. Tobias (Ed.), *Foundations of auditory theory* (Vol. 1). London: Academic Press.
- SHELLENBERG, E. G., & TRAINOR, L. J. (1996). Sensory consonance and the perceptual similarity of complex-tone harmonic intervals: Test of adult and infant listeners. *Journal of the Acoustical Society of America*, **100**, 3321-3328.
- SCHENKER, H. (1979). *Free composition* (E. Oster, Trans.). New York: Longman. (Original work published 1935)
- SCHOENBERG, A. (1978). *Theory of harmony* (R. E. Carter, Trans.). London: Faber and Faber. (Original work published 1911; trans. from 3rd ed., 1922)
- SMITH, B. (1995). PsiExp: An environment for psychoacoustic experimentation using the IRCAM Musical Workstation. In *Proceedings of the Society for Music Perception and Cognition Conference '95* (pp. 83-84). Berkeley, CA.
- TERHARDT, E. (1974). On the perception of periodic sound fluctuation (roughness). *Acustica*, **30**, 201-212.
- TRAMO, M. J. (1996). Neural representations of tonal harmony [Abstract]. *International Journal of Psychology*, **3**, 187.
- VIEMEISTER, N. F. (1977). Temporal factors in audition: A system analy-

- sis approach. In E. F. Evans & J. P. Wilson (Eds.), *Psychophysics and physiology of hearing* (pp. 419-429). London: Academic Press.
- Vos, J. (1986). Purity ratings of tempered fifths and major thirds. *Music Perception*, **3**, 221-258.
- WRIGHT, J. K., & BREGMAN, A. (1987). Auditory stream segregation and the control of dissonance in polyphonic music. *Contemporary Music Review*, **2**, 63-92.
- ZWICKER, E., & FASTL, H. (1990). *Psychoacoustics, facts and models*. Berlin: Springer-Verlag.
- ZWICKER, E., FLOTTORP, G., & STEVENS, S. S. (1957). Critical band width in loudness summation. *Journal of the Acoustical Society of America*, **29**, 548-557.

NOTES

1. The critical band is a general concept defining the frequency resolution of the auditory system and is directly influenced by the biomechanics of the basilar membrane (Scharf, 1970).

2. A factor analysis was not performed after the PCA, for the following reason. A criterion, such as varimax (Kaiser, 1958), rotates the axes to make the subjects' loading on each factor either very high or very low. Such a rotation would alter the consensus among the subjects as to the ordering of the timbres along the first dimension. Our objective was mainly to assess this consensus; therefore, the axes were not rotated.

3. It must be acknowledged that even in the original version of the GTTM (Lerdahl & Jackendoff, 1983, p. 297), timbre was considered as a candidate contributor to nontonal tension perception. However, the salience rules proposed in Lerdahl (1989) were defined for a collection of pitches and not for one sound event that had a homogeneous "timbre." In any case, they did not include roughness as a relevant criterion.

(Manuscript received March 4, 1997;
revision accepted for publication September 18, 1998.)