DETERMINING THE EVEN HARMONIC ATTENUATION AT WHICH THE 'CLARINET-LIKE' TIMBRE OF COMPLEX TONES BECOMES DOMINANT

Ella Manor[1], William L. Martens[2] and Mark Bassett[3]

Faculty of Architecture, Design and Planning
University of Sydney, Sydney NSW 2006, Australia

[1]Email: ella.manor@sydney.edu.au
[2] Email: william.martens@sydney.edu.au
[3] Email: m.bassett@sae.edu

Abstract

It is well known that complex harmonic tones can be varied in their timbral character by adjusting the relative amplitudes of selected sets of harmonics, most obviously exemplified by the reduction of even-harmonic amplitudes to create a 'Clarinet-like' timbre. What is not well known is how great the reduction in even harmonics must be, relative to the odd harmonic amplitudes, in order for the 'Clarinet-like' timbre to be clearly distinguished from the timbre of other complex tones containing energy distributed over all harmonics. For example, a complex waveform with more of an 'Oboe-like' timbre will exhibit an orderly progression of harmonic amplitudes that decrease with increasing frequency. A slight deviation from this progression in which the amplitudes of all even harmonics are reduced as a group, for example by a common attenuation factor of 3dB, does not shift timbre identification responses away from predominantly 'Oboe-like' to predominantly 'Clarinet-like.' It was one motivation for the current research to determine how great the amplitude reduction in the group of even harmonics must be in order for the number of 'Clarinet-like' responses made by a panel of listeners to come to dominate the observed number of 'Oboe-like' responses. A second motivation for the research reported here was to determine how to adjust the relative amplitude of harmonics of a complex tone containing only odd harmonics in order to match the perceived 'sharpness' of a complex tone containing energy distributed more evenly over all harmonics. This was an important question to ask in this investigation because the perceived sharpness of these two complex tones clearly differed even though the rate of spectral roll-off was matched between them. Having the answer to the second question posed above regarding the means by which tonal 'sharpness' can be held constant, a more satisfying answer to the first question is made possible. The answer is more satisfying because it eliminates the confounding influence of changes in the perceived 'sharpness' of timbres occurring within the sequence of complex tones, in which the physical boundary between complex tonal stimuli identified as having 'Oboe-like' versus 'Clarinet-like' timbre was determined. Using a measure termed the 'Even-harmonic Attenuation,' the timbral transition point in categorical response proportions could be associated cleanly with a position on a physical stimulus continuum for such complex tones. This position was determined to be around 9 to 12 dB for the stimuli presented in this study.

1. Introduction

Steady-state complex tones have primary auditory attributes that include pitch, loudness, and timbre. Although psychophysical relations for the first two of these attributes are relatively well understood, psychophysical relations for timbre remain elusive, primarily due to the multidimensional complexity of the psychological phenomena associated with the perception of timbre, but also due to the difficulty of describing effectively the dimensions of the stimulus, even if the dynamic aspects characteristic of musical instrument tones are excluded from consideration. In the earliest known systematic treatment of the timbre of steady-state complex tones, for which a firm experimental base is also given, Helmholtz [1] proposed a number of rules relating physical features to perceptual differences between tones varying in harmonic amplitude pattern. The third rule determines that "*... complex tones consisting of only odd harmonics sound hollow ...",* whereas *"...complex tones with strong harmonics beyond the 6th or 7th sound sharp"* according to the fourth rule [1].

Despite the fact that spectral-temporal variation was excluded from the experimental studies reported by Helmholtz [1], his investigation of timbre did focus upon the simulation of musical tones (and also upon perceived vowel coloration in steady-state vocal timbres). The above two rules apply well to the investigation of steady-state complex tones in the current study, as these complex tones have been intentionally manipulated along the two identified timbral dimensions, one being the growing '*hollowness'* associated with an increasing predominance of odd harmonics, and the other being the growing '*sharpness'* that results from the relative strengthening of higher harmonics. What is novel here is the explicit intent to empirically determine the transition along the '*hollowness'* dimension between an 'Oboe-like' timbre at one end of a series of interpolated tones, and a 'Clarinet-like' timbre at the other end of the series of tones, as the amplitudes of all even harmonics are reduced progressively as a group. Furthermore, care was taken to hold constant the perceived sharpness of each of these interpolated complex tones as they gradually took on more of the 'Clarinet-like' timbre. This required a corrective variation in the rate of spectral roll-off in harmonic amplitude, as complex tones containing only odd harmonics were heard as lower in perceived sharpness than otherwise comparable complex tones containing all harmonics. So two experimental listening tasks were completed in order to provide answers to the following two related questions:

1. When the character of a complex tone is varied through the selective reduction of odd harmonic amplitudes, how can sharpness be held constant?

2. When the perceived sharpness is held constant for each complex tone in a series of tones gradually interpolated between two extreme anchoring stimuli, at what point in the sequence will the tonal stimulus shift from a timbre predominantly identified as 'Oboe-like' to one that is predominantly identified as 'Clarinet-like'?

It seems strange that such a simple study of musical timbre has never been attempting in the proposed manner, as it seems to follow directly from observations made by Helmholtz over 150 years ago. The reason for the omission of such a study from the psychoacoustic literature might be that there was no obvious motivation for seeking an answer to the above two questions. Indeed, the authors would not have pursued the current investigation had it not been for a need to design and generate a set of sound stimuli for a timbral ear-training program, such as that described by Quesnel [2]. It is beyond the scope of this paper to describe the timbral ear training program for which the current set of complex tones have being designed; however, it will suffice to say here that knowledge regarding a timbral transition point along a carefully crafted continuum of sound stimuli is advantageous for training purposes. The reader is directed to the recent paper by McKinnon-Bassett, et al. [3] for more information about this and other ear-training programs.

2. Methods

2.1 Stimuli

The steady-state complex tones explored in this study can be described as a periodic fluctuation of sound pressure *p*, over time *t*, and can be represented by the following equation:

$$p(t) = \sum_{n=1}^{N} a_n \sin(2\pi nft + \phi_n) \qquad (1)$$

It has long been established that the timbre of such steady-state complex tones will primarily depend upon the amplitude pattern a_1, a_2, a_3, …, a_n, and that the timbre depends to a much lesser extent upon the phase pattern, ϕ_1, ϕ_2, ϕ_3, …, ϕ_n of the *N* successive harmonics [4]. The amplitude pattern for a number of canonical waveforms can be described by a sequence of amplitudes that is a simple function of harmonic number, such that the sum of an infinite series of sinusoidal components will converge upon familiar waveform shapes such as that with the shape of a descending ramp; or, with selective deletion of even harmonics, the square-wave. If the spectral amplitude distribution is truncated at a given cut-off frequency (i.e., it is band-limited), then the resulting waveforms will have some visible ripple in comparison to the canonical waveforms. Take for example the Oboe timbre, which exhibits an orderly progression of harmonic amplitudes that decrease with increasing frequency equally for all harmonics (a_n, where *n* =1, 2, 3, 4, 5, etc.) Then consider the Clarinet timbre, which is associated with a waveform that has higher amplitude odd harmonics (a_n, where *n* =1, 3, 5, 7, etc.), and much lower amplitude at even harmonics (a_n, where *n* = 2, 4, 6, 8, etc.) Figure 1 shows the 'Oboe-like' waveform and the 'Clarinet-like' waveform formed from the two sets of harmonics described above. The figure shows a comparison between two different 'Oboe-like' and 'Clarinet-like' waveforms with spectral roll-off factors equal to 4.5 dB/octave (i.e., the standard 'Oboe-like' tone) and 6 dB/octave. The two panels in figure 1 show that as the spectral roll-off increases (i.e., right panel) the shapes of the waveforms are of a descending ramp for the 'Oboe-like' and of a square wave for the 'Clarinet-like'. At 6 dB/octave the waveform shapes match the canonical shapes of square wave and sawtooth; however, as the dB/octave roll-off values increase, the shapes of the waveforms become more sinusoidal. Figure 2 compares between the amplitude roll-off resulting for the standard 'Oboe-like' waveform (i.e., left panel) and the amplitude roll-off resulting for another 'Oboe-like' waveform (i.e., left panel.) The left panels of figures 1 and 2 show the standard 'Oboe-like' waveform plotted in the time domain and in the frequency domain respectively.

For the current experiments, the 'Oboe-like' waveform was produced by progressively summing up constituent harmonics to generate a single cycle of a complex waveform (i.e., containing 32 harmonics) at each of 23 spectral roll-off values, with the fundamental frequency set to 311 Hz (i.e., D#4). The synthetic tone is produced for the set of 23 complex waveforms using a conventional Attack-Decay-Sustain-Release (ADSR) envelope, which attempts to shape the tone's temporal characteristics to those observed in musical instruments. The set of 23 'Clarinet-like' waveforms was similarly produced, however only the odd harmonics were included, by attenuating the amplitudes of the even harmonics to zero (i.e., infinity in dB scale).

The sets of the 'Oboe-like' and 'Clarinet-like' tones were submitted to a spectral centroid analysis, as well as to predicted sharpness analysis [5]. The results of the analyses can be seen in Figure 3. The spectral centroid measured for the standard 'Oboe-like' tone (i.e., having a spectral roll-off of 4.5 dB/octave) may be closely matched to the spectral centroid measured for the 'Clarinet-like' tone, which has a spectral roll-off of 6.2 dB/octave. However, using predicted sharpness measurement, the standard 'Oboe-like' tone can be closely matched to the 'Clarinet-like' tones, which has a spectral roll-off of 8.1 dB/octave. Therefore, one motivation of this study is to find the 'Clarinet-like' tone (from a set of 23 tones) that is closely matched in perceived sharpness to the standard 'Oboe-like' tone. These sets of 'Oboe-like' and 'Clarinet-like' tones were used as stimuli for a sharpness-matching task.

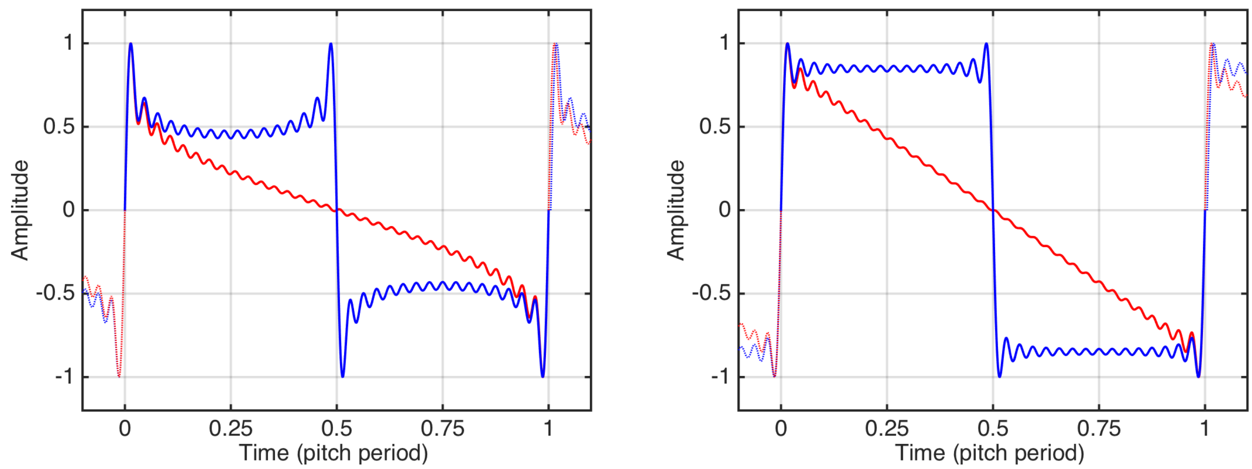

Figure 1. The two panels show a comparison between the waveforms formed from the two sets of harmonics in which the ‘Clarinet-like’ waveform plotted in blue (black for a black and white print) and the ‘Oboe-like’ waveform plotted in red (grey for a black and white print). The left panel shows the resulting ‘Oboe-like’ and ‘Clarinet-like’ waveforms using a spectral roll-off value of 4.5 dB/octave (i.e., set as the standard ‘Oboe-like’ waveform), where the panel on the right shows the resulting waveforms using a spectral roll-off value of 6 dB/octave (resembling the canonical waveshapes).

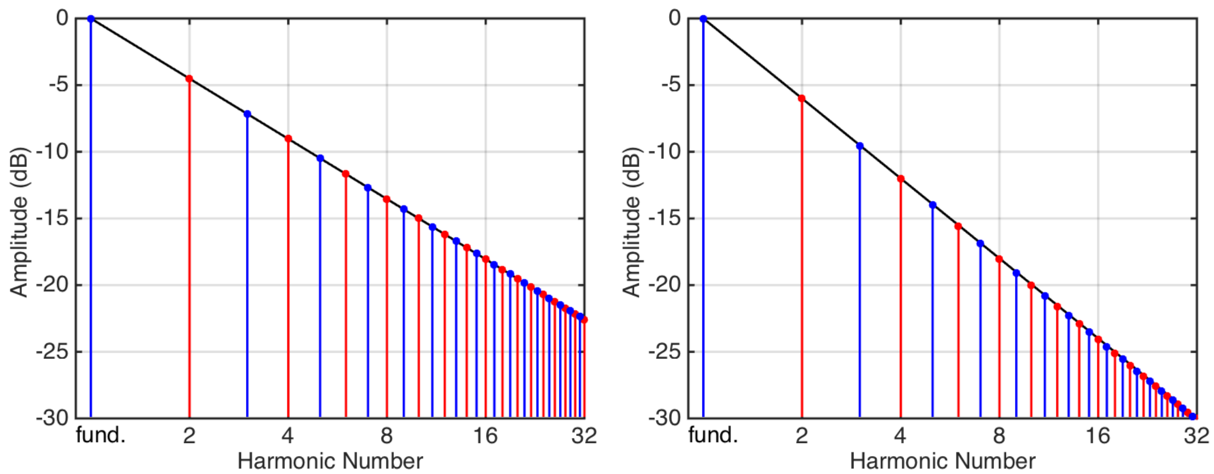

Figure 2. The two panels show plots of the amplitude in dB as a function of harmonic number resulting from applying different spectral roll-off factors for the ‘Oboe-like’ waveform. The amplitude of the fundamental frequency in both cases is set to 0 dB, from which the harmonic amplitudes decrease with increasing frequency in a manner dependent on the spectral roll-off factor. The left and right panels show the plots resulting from applying spectral roll-off factors of 4.5 dB/octave and 6 dB/octave respectively. The blue lines (black for a black and white print) indicate the odd harmonic amplitudes, whereas the red lines (grey for a black and white print) represent the even harmonic amplitudes.

The tones used in a subsequent categorical-identification task were produced based upon the result of the sharpness-matching task, which was designed to produce a close match between the ‘Oboe-like’ and the ‘Clarinet-like’ tones in terms of perceived sharpness. The selected ‘Clarinet-like’ tone was then modified to produce a set of six stimuli differing in their even-to-odd harmonic amplitude ratio. Each of the ‘Clarinet-like’ tones were produced in the same manner as described above, however instead of attenuating the amplitude of the even harmonics to zero, the amplitudes were multiplied by a scaling factor specified in decibels. Six factors were used to produce each of the stimuli; -3, -6, -9, -12, -15, and -21 dB, where the tones produced at each of the extreme dB factors are more similar to either the standard ‘Oboe-like’ tone or the matched ‘Clarinet-like’ tone.

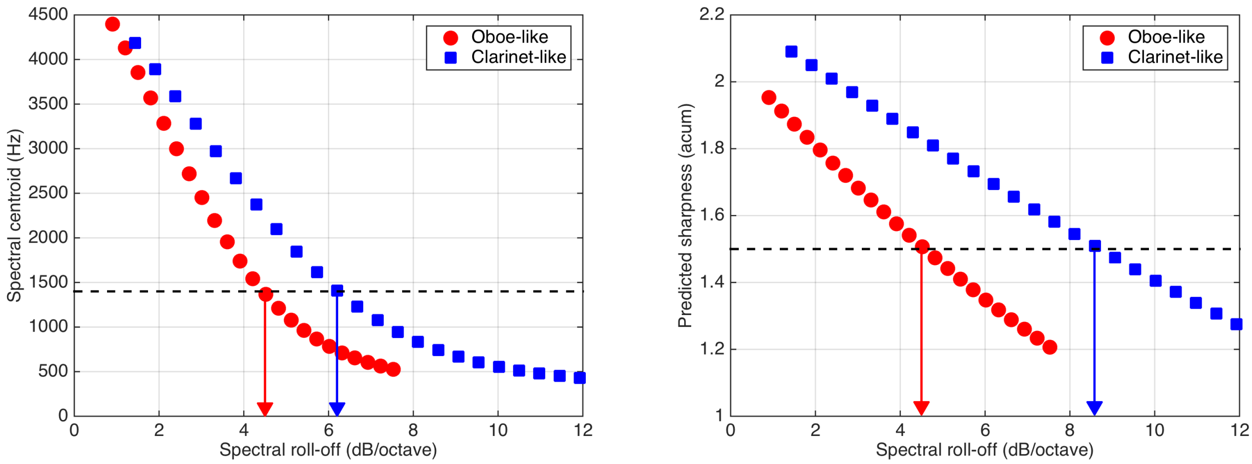

Figure 3. The left panel shows the spectral centroid measured in Hertz for each of the 23 ‘Oboe-like’ tones (red circles) and for the 23 ‘Clarinet-like’ tones (blue squares) plotted as functions of spectral roll-off. The right panel shows predicted sharpness measured in acum for each of the 23 ‘Oboe-like’ tones (red circles) and for the 23 ‘Clarinet-like’ tones (blue squares) plotted as functions of spectral roll-off. The red arrow indicates the ‘Oboe-like’ standard tone and the blue arrow indicates the closely matched ‘Clarinet-like’ tone according to the measurements.

2.2 Procedure

A total of 21 listeners participated, with 10 participating in both tasks out of which seven listeners performed the categorical-identification task for a single run, and three listeners completed ten such consecutive runs, in which data combined into a single result from each listener. The two tasks included sharpness matching and categorical-identification. Listeners were given an oral orientation prior to the sharpness-matching task, which included a detailed description of the assessment types and the timbral attribute tested (i.e., perceived sharpness), and they were also provided with an instruction sheet for operating the graphical user interface (GUI).

The sharpness-matching task included two types of assessments, and conducted in successive listening tests, where the first assessment can be defined as “apples-to-apples” comparison and the second assessment was more of “apples-to-oranges” comparison. First, listeners were asked to complete a matching task five times, with the perceived sharpness of a variable comparison ‘Oboe-like’ tone (i.e., taken from the 23 tones differing in spectral roll-off) heard just after a presentation of a fixed standard ‘Oboe-like’ tone that had a spectral roll-off of 4.5 dB/octave. Then, in a second task, listeners were asked to match the perceived sharpness of a variable comparison ‘Clarinet-like’ tone (i.e., one from the series of 23 tones) with the standard ‘Oboe-like’ tone. Data collection was conducted using a separate GUI for each of the tasks, where by pressing the button labelled ‘A’, the standard ‘Oboe-like’ stimulus could be played, and the comparison stimulus could be played by pressing the button labelled ‘B’. Listeners were asked to adjust the perceived sharpness of ‘B’ by clicking the ‘Sharper’ or ‘Darker’ response buttons. If ‘B’ were heard as darker than ‘A’, pushing the ‘Sharper’ button would make stimulus ‘B’ sharper on its next presentation. When listeners were satisfied with the sharpness match between the two tones, they were to continue to the next trial by clicking-on the ‘Next’ button displayed in the GUI.

Based upon the results of the sharpness-matching task, described below, a set of six tones, as well as the standard ‘Oboe-like’ tone and a sharpness-matched ‘Clarinet-like’ tone were selected as stimuli for the categorical-identification task. Then the categorical-identification task listeners were presented consecutively with three different stimuli; the standard ‘Oboe-like’ tone (i.e., labelled as A), the matched ‘Clarinet-like’ tone (i.e., labelled as B), and a randomly selected tone from the set of six interpolated tones (i.e., labelled as X). Listeners were asked to indicate whether tone X sounded more similar to tone A or to tone B. The six tones for matching (i.e., always labelled as X) were presented in a random order for a total of five double-blind runs of matching.

3. Results

The results of the sharpness-matching task, for both comparisons within timbral type (only 'Oboe-like' tones) and between timbral types ('Oboe-like' versus 'Clarinet-like') are shown in Figure 4. The left panel of the figure plots the proportion of 'sharper' responses observed when a fixed standard stimulus was compared with a variable comparison stimulus that differed only in terms of spectral roll-off. Therefore, when spectral roll-off was matched between the standard and comparison stimuli, the comparison was between identical waveforms. As the spectral roll-off value for the standard stimulus was held constant across trials at 4.5 dB/octave, this was the point at which the proportion of 'sharper' responses passed through the point of subjective equality operationally defined as the 0.5 point on the curve. This task was completed first by all subjects, partly for sake of training on a relatively easy task, in which an "apples-to-apples" comparison was required. Then, in a subsequent test, listeners were asked to match the perceived sharpness of 'Clarinet-like' tones of varying spectral roll-off to a fixed roll-off standard 'Oboe-like' tone. The right panel of Figure 4 plots the proportion of 'sharper' responses observed in this latter case, showing that a 'Clarinet-like' tone needed to have spectral roll-off of around 6.2 dB/octave to match its sharpness to an 'Oboe-like' tone with spectral roll-off fixed at 4.5 dB/octave.

The proportion data resulting from the two sharpness-matching tasks were transformed to z-scores in order to allow for a linear regression analysis to be performed on the originally curvilinear proportion data. The resulting regression lines are illustrated in Figure 5. The point of subjective equality for sharpness in the two examined cases could be determined from the zero-crossing point of the lines fit to the z-scores, with results that agreed with that based upon the smooth plotted curves shown in Figure 4. These results of the sharpness-matching task established the anchoring stimuli at either end of a continuum of similar tones, which enabled the construction a set of interpolated tones for the subsequent categorical-identification task. The two anchoring stimuli, one an 'Oboe-like' tone and the other a sharpness-matched 'Clarinet-like' tone, were presented on every trial along with a third tone that could sound more or less like each of the anchors. According to a stimulus order that was determined through random sampling (without replacement), six such interpolated tones were judged as more similar either to the 'Oboe-like' tone or the 'Clarinet-like' tone.

Figure 6 shows the results of the categorical-identification task for three individual listeners (labelled S1, S2, and S3), and for the combined data from ten listeners (labelled All Subjects). In all four plots, the x-axis tick marks indicate the six different values of the measure termed 'Even-harmonic Attenuation' that were selected for the experimental stimuli that were compared to the two anchoring tones. The greatest amount of attenuation applied selectively to even-harmonic amplitudes was 21 dB, which produced tones with timbre quite similar to the 'Clarinet-like' anchoring tone. For this interpolated tone at such extreme even-harmonic attenuation, the tone was almost always identified as 'Clarinet-like,' with observed response proportions typically above 0.9, and above 0.8 for all listeners who participated. The mildest amount of attenuation that was applied to even-harmonic amplitudes was 3 dB, which produced tones with timbre quite similar to the 'Oboe-like' tone. The associated 'Clarinet-like' response proportions were typically below 0.1, and were always below 0.2 for all listeners who participated.

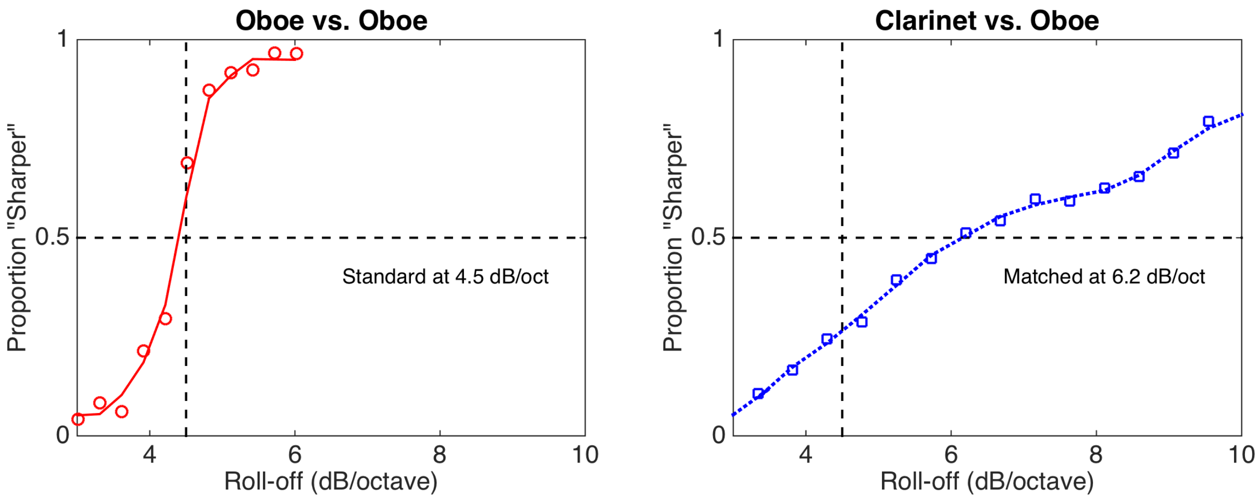

Figure 4. The results of the sharpness-matching task for comparisons <u>within</u> timbral type (left panel – for only 'Oboe-like' tones), and for comparisons <u>between</u> timbral types (right panel – for 'Oboe-like' versus 'Clarinet-like' tones).

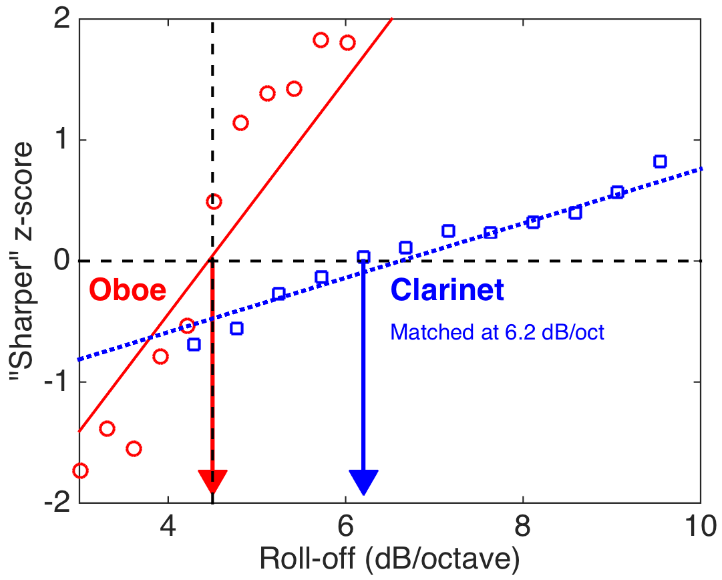

Figure 5. Regression line analysis of the results of the two sharpness-matching tasks, using z-score transformed proportion data for comparisons <u>within</u> timbral type versus comparisons <u>between</u> timbral types (red circular symbols show the z-score values for matching 'Oboe-like' to 'Oboe-like' tones, and blue square symbols show the z-score values for matching 'Oboe-like' to 'Clarinet-like' tones).

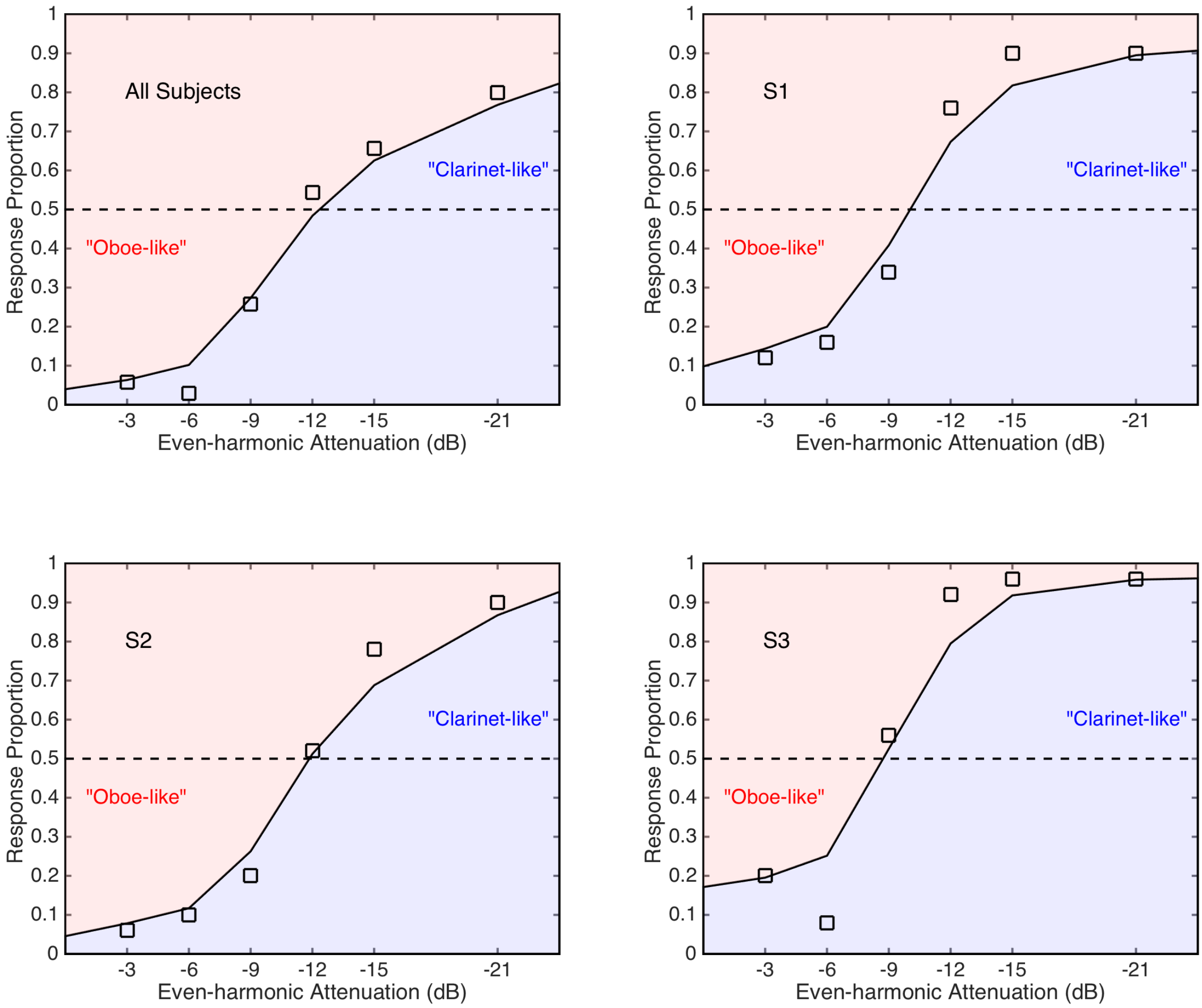

Figure 6. The results of the categorical-identification task are shown in separate plots for each of three individual listeners (labelled S1, S2, and S3), and for the combined data from ten listeners (labelled All Subjects). The x-axis tick marks indicate the dB values of the experimental stimuli in terms of the measure termed 'Even-harmonic Attenuation.' The y-axis values show the 'Clarinet-like' response proportions for each of the six tones positioned along a continuum between 'Oboe-like' and 'Clarinet-like' anchoring tones.

4. Discussion and Conclusion

The integrated results of two timbre comparison tasks for complex tones having 'Oboe-like' versus 'Clarinet-like' timbres enable a strong conclusion regarding the following question:

> As the amplitudes of even-harmonic components are progressively reduced to transform an 'Oboe-like' tone to one that is heard as 'Clarinet-like,' at what point in the progression will the tonal stimulus shift from a timbre predominantly identified as 'Oboe-like' tone to one that is predominantly identified as 'Clarinet-like?'

In order to avoid the influence of changing tonal sharpness on this categorical-identification, an attempt was made to hold the perceived sharpness constant for each complex tone in a series of tones gradually interpolated between two extreme anchoring stimuli. This required completion of a sharpness-matching experiment to determine how to adjust the amount of spectral roll-off in the 'Clarinet-like' tone so that it would match the 'Oboe-like' tone in terms of perceived sharpness.

The results of the sharpness-matching task in two separate cases showed that comparing "apples-to-apples" (juxtaposing two 'Oboe-like' tones) produced a steeper transition in response proportions than did the "apples-to-oranges" comparison (in which 'Oboe-like' tones were juxtaposed with 'Clarinet-like' tones). The answer to the first question, at least in the case of the present study, was that a 'Clarinet-like' tone needed to have spectral roll-off of around 6.2 dB/octave to match its sharpness to an 'Oboe-like' tone with spectral roll-off fixed at 4.5 dB/octave.

Having answered this prior question, regarding the means by which tonal 'sharpness' can be held constant, a more satisfying answer to the more central question was made possible. By varying even-harmonic attenuation along a continuum of sharpness-matched tones, the timbral transition point in categorical response proportions could be associated cleanly with a position on a physical stimulus continuum for such complex tones, and this position was determined to be around 9 to 12 dB for the stimuli presented in this study.

References

[1] Helmholtz, H.L.F. *Die Lehre von den Tonempfindungen als Physiologische Grundlage für die Theorie der Musik*. Germany: von Friedrich Vieweg und sohn, 1863.

[2] Quesnel, R. "Timbral ear trainer: Adaptive, interactive training of listening skills for the evaluation of timbre differences", *Proceedings of the 100th Convention of the Audio Engineering Society*, Copenhagen, Denmark, May 1996.

[3] McKinnon-Bassett, M., Martens, W. and Cabrera, D. "Experimental comparison of two versions of a technical ear training program: Transfer of training on tone color identification to a dissimilarity-rating task", *Proceedings of the Audio Engineering Society Conference 50th International Conference: Audio Education*, Murfreesboro, USA, 25-27 July 2013.

[4] Plomp, R. and Steeneken, H. J. M. "Effect of phase on the timbre of complex tones", *Journal of the Acoustical Society of America,* **46**, 409-421, (1969).

[5] Zwicker, E. and Fastl, H. *Psychoacoustics: Facts and Models*, second edition, Springer-Verlag, 1990.